

```

1 ATGGAGCGCA CAGCAGGCAA AGAGCTGGCC CTGGCACCGC TGCAGGACTG
51 GGGTGAAGAG ACCGAGGACG GCGCGGTGTA CAGTGTCTCC CTGCGGCGGC
101 AGCGCAGTCA GCGCAGGAGC CCGCGGAGG GCGCGGGGGG CAGCCAGGCT
151 CCCAGCCCCA TTGCCAATAC CTTCCTCCAC TATCGAACCA GCAAGGTGAG
201 GGTGCTGAGG GCAGCGCGCC TGGAGCGGCT GGTGGGAGAG TTGGTGTGTTG
251 GAGACCGTGA GCAGGACCCC AGCTTCATGC CCGCCTTCCT GGCCACCTAC
301 CGGACCTTTG TACCCACTGC CTGCCTGCTG GGCTTTCTGC TGCCACCAAT
351 GCCACCGCCC CCACCTCCCG GGGTAGAGAT CAAGAAGACA GCGGTACAAG
401 ATCTGAGCTT CAACAAGAAC CTGAGGGCTG TGGTGTCACT GCTGGGCTCC
451 TGGCTGCAGG ACCACCCTCA GGATTTCCGA GACCACCCTG TCCATTCCGA
501 CCTGGGCAGT GTCGAACCT TTCTGGGCTG GGCGGCCCCA GGGAGTGCTG
551 AGGCTCAAAA AGCAGAGAAG CTTCTGGAAG ATTTTTTGA GGAGGCTGAG
601 CGAGAGCAGG AAGAGGAGCC GCCTCAGGTG TGGACAGGAC CTCCCAGAGT
651 TGCCCAAAC TCTGACCCAG ACTCTTCAGA GGCCTGCGCG GAGGAAGAGG
701 AAGGGCTCAT GCCTCAAGGT CCCAGCTCC TGGACTTCAG CGTGGACGAG
751 GTGGCCGAGC AGCTGACCCT CATAGACTTG GAGCTCTTCT CCAAGGTGAG
801 GCTCTACGAG TGCTTGGGCT CCGTGTGGTC GCAGAGGGAC CGGCCGGGGG
851 CTGCAGGCGC CTCCCCACT GTGCGCGCCA CCGTGGCCA GTTCAACACC
901 GTGACCGGCT GTGTGCTGGG TTCCGTGCTC TGAGCACCGG GCTTGCGCGC
951 CCCGAGAGG GCGCAGCGGC TGGAGAAGTG GATCCGCATC GCCCAGCGCT
1001 GCCGAGAACT GCGGAACCTC TCCTCCTTGC GCGCCATCCT GTCCGCCCTG
1051 CAATCTAACC CCATCTACCG GCTCAAGCGC AGCTGGGGGG CAGTGAGCCG
1101 GGAACCGCTA TCTACTTTCA GGAACTTTC GCAGATTTTC TCCGATGAGA
1151 ACAACCACT CAGCAGCAGA GAGATTCTTT TCCAGGAGGA GGCCACTGAG
1201 GGATCCCAAG AAGAGGACAA CACCCAGGC AGCCTGCCCT CAAAACCACC
1251 CCCAGGCCCT GTCCCCTACC TTGGCACCTT CTTACGGAC CTGTTATGTC
1301 TGGACACAGC CCTGCCGGAT ATGTTGGAGG GGGATCTCAT TAACTTTGAG
1351 AAGAGGAGGA AGGAGTGGGA GATCCTGGCC CGCATCCAGC AGCTGCAGAG
1401 GCGCTGTCA AGCTACACCC TGAGCCCCA CCGGCCATC CTGGCTGCCC
1451 TGATGCCCCA GAACAGCTC ACCGAGGAGC AGAGTACCG GCTCTCCCGG
1501 GTCATTGAGC CACCAGCTGC CTCCTGCCCC AGCTCCCCAC GCATCCGACG
1551 GCGGATCAGC CTCACCAAGC GTCTCAGTGC GAAGCTTGCC CGAGAGAAAA
1601 GCTCATCACC TAGTGGGAGT CCCGGGGACC CCTCATCCCC CACCTCCAGT
1651 GTGTCCCAG GGTACCCCC CTCAAGTCCT AGAAGCAGAG ATGCTCCTGC
1701 TGGCAGTCCC CCGGCCTCTC CAGGGCCCCA GGGCCCCAGC ACCAAGCTGC
1751 CCCTGAGCCT GGACCTGCCC AGCCCCCGG CTTTCGCTTT GCCTCTGGGC
1801 AGCCCTCGAA TCCCCCTCCC GGCGCAGCAG AGCTCGGAGG CCCGTGTCAT
1851 CCGCGTCAGC ATCGACAATG ACCACGGGAA CCTGTATCGA AGCATCTTGC
1901 TGACCAGTCA GGACAAAGCC CCCAGCGTGG TCCGGCGAGC CTTGCAGAAG
1951 CACAATGTGC CCCAGCCCTG GGCCTGTGAC TATCAGTCT TTCAAGTCCT
2001 TCCTGGGGAC CGGGTGCTCC TGATTCTTGA CAATGCCAAC GTCTTCTATG
2051 CCATGAGTCC AGTCGCCCC AGAGACTTCA TGCTGCGGCG GAAAGAGGGG
2101 ACCCGGAACA CTCTGTCTGT CTCCCCAAGC TGA (SEQ ID NO:1)

```

FEATURES:

Start Codon: 1

Stop Codon: 2131

FIGURE 1A

Homologous proteins:
Top 10 BLAST Hits

	Score	E
CRA 160000003781983 /altid=gi 8650435 /def=gb AAF78208.1 AF2376...	1119	0.0
CRA 103000001517213 /altid=gi 7682471 /def=gb AAF67280.1 AF1867...	448	e-125
CRA 18000005222968 /altid=gi 4589562 /def=dbj BAA76803.1 (AB02...	448	e-125
CRA 38000046964621 /altid=gi 10185686 /def=gb AAG14400.1 AF1867...	448	e-125
CRA 18000004932640 /altid=gi 8394180 /def=ref NP_058542.1 ral ...	442	e-123
CRA 18000005115144 /altid=gi 4758532 /def=ref NP_004752.1 RAB2...	387	e-106
CRA 120000015982361 /altid=gi 9931302 /def=gb AAG02122.1 AF2957...	385	e-105
CRA 18000004933686 /altid=gi 6677735 /def=ref NP_033084.1 ral ...	372	e-102
CRA 18000005176024 /altid=gi 3811378 /def=gb AAC69894.1 (AF100...	369	e-101
CRA 18000005033177 /altid=gi 6677737 /def=ref NP_033085.1 ral ...	368	e-100

BLAST dbEST hits:

	Score	E
gi 12067555 /dataset=dbest /taxon=96...	880	0.0
gi 11311915 /dataset=dbest /taxon=96...	767	0.0
gi 12067611 /dataset=dbest /taxon=96...	700	0.0
gi 843816 /dataset=dbest /taxon=9606 /...	597	e-168
gi 3331474 /dataset=dbest /taxon=9606 ...	591	e-166
gi 6660347 /dataset=dbest /taxon=9606 ...	460	e-127
gi 2322877 /dataset=dbest /taxon=9606 ...	424	e-116
gi 2348656 /dataset=dbest /taxon=9606 ...	422	e-115
gi 4189370 /dataset=dbest /taxon=9606 ...	353	1e-94

EXPRESSION INFORMATION FOR MODULATORY USE:

Library source:

gi|12067555 bocio-tumor
gi|11311915 Head-neck
gi|12067611 bocio-tumor
gi|843816 Breast
gi|3331474 Parathyroid tumor
gi|6660347 Uterus papillary carcinoma
gi|2322877 Colon tumor
gi|2348656 Colon tumor
gi|4189370 Pooled human melanocyte, fetal heart, and pregnant uterus

FIGURE 1B

```

1 MERTAGKELA LAPLQDWGEE TEDGAVYSVS LRRQRSQRRS PAEGPGGSQA
51 PSPIANTFLH YRTSKVRVLR AARLERLVGE LVFGDREQDP SFMPAFLATY
101 RTFVPTACLL GFLPPMPPP PPPGVEIKKT AVQDLSFNKN LRAVSVLGS
151 WLQDHPQDFR DHPVHSDLGS VRTFLGWAAP GSAEAQKAEK LLEDLEEAE
201 REQEEPPQV WTGPPRVAQT SDPDSSEACA EEEGLMPQG PQLDFSVD
251 VAEQLTLIDL ELFSKVRLYE CLGSVWSQRD RPGAAGASPT VRATVAQFNT
301 VTGCVLGSVL GAPGLAAPQR AQRLEKWIRI AQRCRELRNF SSLRAILSAL
351 QSNPIYRLKR SWGAVSREPL STFRKLSQIF SDENNHLSR EILFQEEATE
401 GSQEDNTPG SLPSKPPPGP VPYLGTF LTD LVMLDTALPD MLEGDLINFE
451 KRRKEWEILA RIQQLQRRQ SYTLSPHPPI LAALHAQNL TEEQSYRLSR
501 VIEPPAASCP SSPRIIRRRIS LTKRLSAKLA REKSSSPSGS PGDPSSPTSS
551 VSPGSPSSP RSRDAPAGSP PASPGPGGPS TKLPLSLDLP SPRPFALPLG
601 SPRIPPLPAQQ SSEARVIRVS IDNDHGNLYR SILLTSQDKA PSWRRALQK
651 HNVPQWACD YQLFQVLPD RVLLIPDNAN VFYAMSPVAP RDFMLRRKEG
701 TRNTLSVSPS (SEQ ID NO:2)

```

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

339-342 NFSS

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
CAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 3

1	374-377	RKLS
2	517-520	RRIS
3	523-526	KRLS

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 18

1	30-32	SLR
2	342-344	SLR
3	36-38	SQR
4	277-279	SQR
5	63-65	TSK
6	99-101	TYR
7	170-172	SVR
8	36-38	SQR
9	277-279	SQR
10	290-292	TVR
11	30-32	SLR
12	342-344	SLR
13	372-374	TFR
14	388-390	SSR
15	495-497	SYR
16	512-514	SPR
17	559-561	SPR
18	591-593	SPR

FIGURE 2A

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 9

1	40-43	SPAE
2	221-224	SDPD
3	247-250	SVDE
4	256-259	TLID
5	277-280	SQRD
6	388-391	SSRE
7	402-405	SQEE
8	540-543	SPGD
9	635-638	TSQD

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 13

1	24-29	GAVYSV
2	169-174	GSVRTF
3	181-186	GSAEAQ
4	273-278	GSWSQ
5	283-288	GAAGAS
6	286-291	GASPTV
7	303-308	GCVLGS
8	307-312	GSVLGA
9	410-415	GSLPSK
10	425-430	GTFLTD
11	554-559	GSPPSS
12	568-573	GSPPAS
13	700-705	GTRNTL

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	102	122	1.600	Certain
2	298	318	1.509	Certain
3	416	436	0.748	Putative

FIGURE 2B

BLAST Alignment to Top Hits:

CRA|160000003781983 /altid=gi|8650435 /def=gb|AAF78208.1|AF237669_1
 (AF237669) RalGDS-like protein 3 [Mus musculus] /org=Mus
 musculus /taxon=10090 /dataset=nraa /length=709
 Length = 709

Score = 1119 bits (2862), Expect = 0.0

Identities = 570/714 (79%), Positives = 623/714 (86%), Gaps = 9/714 (1%)

Query: 1 MERTAGKELALAPLQDWGEETEDGAVYSVSLRRQRSQRRSPAEGPGGSQAPSPANTFLH 60
 MERTAGKELALAPLQDWGEETEDGAVYSVSLRRQRSQR +P E G Q P P +TFLH
 Sbjct: 1 MERTAGKELALAPLQDWGEETEDGAVYSVSLRRQRSQRSTP-ERSGEGQTPIPATDTFLH 59

Query: 61 YRTSKVRVLRAARLERLVGELVFGDREQDPSFMPAFLATYRTFVPTACLLGFLPPMPPP 120
 YRTSKVR LRAARLERLV ELV GDREQDP F+PAFLAT+R FVPTA +LGFLPP PPP
 Sbjct: 60 YRTSKVRALRAARLERLVHELVSQDREQDPGFVPAFLATHRAFVPTARVLGFLPPPPPP 119

Query: 121 PPP--GVEIKKTAVQDLSFNKNLRAWSVLGSWLQDHPQDFRDHPVHSDLGSVRTFLGWA 178
 PPP GV+ K+T QDL+F+KNLRAWSVLGSWL++HPQDFRD P H +LG+VR FLGW
 Sbjct: 120 PPPPAGVDSKRTEGQDLNFSKNLRAWSVLGSWLNRNHPQDFRDPPDHQNLGNVRIFLGW 179

Query: 179 APGSAEAQKAELLEDFLAEAEREQ-EEEPQWGTGPPRVAQTSDDPSSEACAEEEGLM 237
 APG AEA++AEKLLDFL+EA+ EQ EEE W+GPPR+AQT + +E C EEE G
 Sbjct: 180 APGAEAREAEKLLDFLKEAKGEQTEEEKRLAWSGPPRIAQTPGSEFAEDCVEEE-GPS 238

Query: 238 PQGPQLLDFSVDVAEQLTLIDLELFSKVRLYECLGSWWSQRDRPGAAGASPTVRATVAQ 297
 +GP+LLDFSVD+VAEQLTL+D+ELF +VR ECLGS+WSQRDRPGAAG SPTVRATVAQ
 Sbjct: 239 SEGPELLDFSVDVAEQLTMDVELFLVRSCCLGSMWSQRDRPGAAGISPTVRATVAQ 298

Query: 298 FNTVTGCVLGSVLGAPGLAAPQRAQRLEKWIRIAQRCRELNFSSLRAILSALQSNPIYR 357
 FNTVTGCVLGSVL APGLAA QRAQR+EKWIRIAQRCRELNFSSLRAILSALQSNPIYR
 Sbjct: 299 FNTVTGCVLGSVLAAPGLAASQRAQRLEKWIRIAQRCRELNFSSLRAILSALQSNPIYR 358

Query: 358 LKRSWGAVSREPLSTFRKLSQIFSDENNHLSSREILFQEEATEGSQEEDNTPGSLPSKPP 417
 LKRSWGAVSREPLS FRKLSQIFSDENHLSSR IL QEE TE ++D GSLPSK P
 Sbjct: 359 LKRSWGAVSREPLSVFRKLSQIFSDENNHLSSRAILSQEETTE---DDDCPSGSLPSKLP 415

Query: 418 PGPVPYLGTFITDLVMLDTALPDMLEGDLINFEKRRKEWEILARIQQLQRCQSYTLSPH 477
 PGPVPYLGTFITDLVMLDTALPD L+G+LINFEKRRKEWEILARIQQLQ+RCQ Y+LSP
 Sbjct: 416 PGPVPYLGTFITDLVMLDTALPDTLKGNLINFEKRRKEWEILARIQQLQRCQRYSLSPR 475

Query: 478 PPILAALHAQNQLTEEQSYRLSRVIEPPAASCPSSPRIRRRISLTKRLSAKLAREKSSSP 537
 PPILAAL AQ QL+EEQSYR+SRVIEPPAASCPSSPRIRRRISLTKRLSAKL+REK+SSP
 Sbjct: 476 PPILAALRAQRQLSEEQSYRVSRIEPPAASCPSSPRIRRRISLTKRLSAKLSREKNSSP 535

Query: 538 SGSPGDPSSPTSSVSPGSPSSPRSRD-APAGSPASPQPGPSTKLPLSLDLPSRPFA 596
 GSPGDPSSPTSSVSPGSPSSPR+R+ P GSPPASPQPGP PSTKL L++D P P P
 Sbjct: 536 GSPGDPSSPTSSVSPGSPSSPRNREPPPGSPASPQPGPSTKLSLTMDPPGPWPVT 595

Query: 597 LPLGSPRIPLPAQSSSEARVIRVSIIDNDHGNLYRSILLTSQDKAPSVVRRALQKHNVQP 656
 L S R+PL QQ+SEARVIRVSI+H+HGNLYRSILLT QDKAPSVV+RAL+KHNVQP
 Sbjct: 596 LTPSSSRVPLLGGQTSEARVIRVSIINNHNHGNLYRSILLTCQDKAPSVVQRALEKHNVQP 655

Query: 657 WACDYQLFQVLPQDRVLLIPDNANVFYAMSPVAPRDFMLRRKEGTRNTLSVSPS 710
 WA DYQLFQVLPQDR LLIPD ANVFYAMSP AP DF+LRRKEGT +TLS SP+
 Sbjct: 656 WARDYQLFQVLPQDRELLIPDGANVFYAMSPAAPGDFLLRRKEGTGHTLSASPT 709 (SEQ ID NO:4)

FIGURE 2C

>CRA|103000001517213 /altid=gi|7682471 /def=gb|AAF67280.1|AF186779_1
(AF186779) RGL protein [Homo sapiens] /org=Homo sapiens
/taxon=9606 /dataset=nraa /length=768
Length = 768

Score = 448 bits (1141), Expect = e-125
Identities = 291/744 (39%), Positives = 409/744 (54%), Gaps = 75/744 (10%)

Query: 11 LAPLQDWGEETEDGAVYSVSLRRQRSQRRSPAE----GPGGSQAPSPIANTFLHYRTSKV 66
++ +QDWGEE E+GAVY V+L+R + Q+ + G G Q P +T Y T K+
Sbjct: 9 MSSIQDWGEEVEEGAVYHVTLLKRVQIQQAANKGARWLGVGDQLPP--GHTVSQYETCKI 66

Query: 67 RVLRAARLERLVGELVFGDREQPSFMPAFLATYRTFVPTACLLGFLPPMPPPPPGVE 126
R ++A LE+LV L+ + D +++ FL+TYR F T +L LL P E
Sbjct: 67 RTIKAGTLEKLVENLLTAFGNDFTYISIFLSTYRGFASTKEVLELLLDRYGNLTSPNCE 126

Query: 127 IKKTAVQDLSFNKNLRAVSVLGSWLQDHPQDFRDHPVHSDLGSVRTFLGWAAPGSAAEQ 186
+ S A+ S+L +WL +DFR+ P L + +L PGS +
Sbjct: 127 EDGSQSSSESKMIRNAIASILRAWLDQCAEDFREPPHFPCLQKLLDYLTRMMPGSDPER 186

Query: 187 KAEKLLDFLEEAEREQEEPPQWGTGPPRVAQTSDDPSSEACAEEEGLMPQGPQLLDF 246
+A+ LLE F ++QE E 'G P S EEEE + + F
Sbjct: 187 RAQNLEQF-----QKQEVETDN---GLPNTISFS-----LEEEEEELEGGESAEFTCF 231

Query: 247 SVDEVAEQLTLIDLELFSKVRLYECLGSVWSQRDRPGAAGASPTVRATVAQFNTVTGCVL 306
S D VAEQLT +D +LF KV + CLG +WS+RD+ +PT+RAT++QFNT+T CV+
Sbjct: 232 SEDLVAEQLTYMDAQLFKKWPHHCLGCIWSRRDKKENKHLAPTIRATISQFNTLTKCV 291

Query: 307 GSVLGAPGLAAPQRAQRLEKWIRIAQRCRELNFSSSLRAITLSALQSNPTIYRLKRSWGAVS 366
++LG L QRA+ +EKWI IA CR L+NFSSSLRAIT+SALQSN IYRLK++W AV
Sbjct: 292 STILGGKELKTQQRAKIIEKWINTIAHECRLLKNFSSSLRAITVSALQSNSTIYRLKKTWAAVP 351

Query: 367 REPLSTFRKLSQIFSDENNHLSSREILFQEEATEGS-----QEEDNTPGSLPSKPP-- 417
R+ + F +LS IFSD NNHL+SRE+L +E ++ + + + T L +
Sbjct: 352 RDRMLMFEELSDIFSDHNNHLTSRELLMKEGTSKFANLDSSSVKENQKRTQRRLQLQKDMG 411

Query: 418 --PGVPYLGTFITDLVMLDTALPDMLEGDLINFEKRRKEWEILARIQQLQRRCSYTLS 475
G VPYLGTFITDL MLDTAL D +EG LINFEKRR+E+E++A+I+ LQ C SY ++
Sbjct: 412 VMQGTVPYLGTFITDLTMDLTALQDYIEGGLINFEKRRREFEVIAQIKLLQSACNSYOMT 471

Query: 476 PHPPILAALHAQNQLTEEQSYRLSRVIEPPAASCPSSPRIRRRISLTKRLSAK-LAREKS 534
P + Q LTEE+SY LS IE A + +SP+ R+ S+ KRLS L +
Sbjct: 472 PDQKFQWFORQQLLTEESYALSCEIEAAADASTTSPKPRK--SMVKRLSLLFLGSDMI 529

Query: 535 SSPSGSPGDPSSPTSSVSPGSPPS-----SPRSRDAPAGSPASPQPQGPSTKLP----- 584
+SP+ + P S S S S S +A GS P P KL
Sbjct: 530 TSPTPTKEQPKSTASGSSGESMDSVSVSSCESNHSEAEEGSITPMDTPDEPQKKLSESSS 589

Query: 585 -----LSLDLPSRPFAL--PLGSPRI-----PLPAQQSSEA 614
S+D S +L PL SP P+ QQ+ +
Sbjct: 590 SCCSIHSDMTNSSGMSSLINPLSSPPSCNNNPKIHKRSVSVTSITSTVLPPVYNQONEDT 649

Query: 615 RVIRVSIDNDHGNLYRSILLTSQDKAPSVRRALQKHNVQPWACDYQLFQVLPGDRVLL 674
+IR+S++++GN+Y+SI+LTSQDK P+V++RA+ KHN+ A +Y+L QV+ D+ L+
Sbjct: 650 CIIRISVEDNNGNMYKSIMLTSQDKTPAVIQRAMLKHNLDSDPAEEYELVQVISEDKELV 709

FIGURE 2D

Query: 675 IPDNANVFYAMSPVAPRDFMLRRK 698
IPD+ANVFYAM+ DF+LR+K
Sbjct: 710 IPDSANVFYAMNSQVNFDFILRKK 733 (SEQ ID NO:5)

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00617	RasGEF domain	226.4	4.3e-64	1
PF00788	Ras association (RalGDS/AF-6) domain	60.6	3.8e-15	1
PF00618	Guanine nucleotide exchange factor for Ras-1	12.0	0.079	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00618	1/1	63	113 ..	1	60 [.	12.0	0.079
PF00617	1/1	244	454 ..	1	227 []	226.4	4.3e-64
PF00788	1/1	613	700 ..	1	112 []	60.6	3.8e-15

FIGURE 2E

```

1  GTCCCCGCC CACCTCCCTG GGAAGCTCT CACTCCCCAA GGAAGCCCAA
51  GACGTCAGAG ACCCTGTCCC GTCTCAAGCT CCGCCTCCAA GGGAAAGAAG
101 TCCCACCCTG TCCCCGAAAA GCTCGGAGAC CTATCTGCAG TTCTGGAGCA
151 GCCCGAGACA GAAGACCTAT CCTGCCCGCA GCAGAGGGCC ATTTCTTCAG
201 GAGGCCATCC CACAATGGCC TCCCTCGAAT CTTTACTGTT CCCAAGAGGC
251 CTCCCCGAAA CCTTTACATA TAGAGTTAGA GTTCAGAGGC TCCCCATCA
301 CAGACGTGGC CTTGACCTAA GTCCCCGCCC CCAATGGCTA GGATCGACCC
351 TGCAGGTGGG CCCTTACCTC CTAGGCCCCG CCCCCGAAAG AAAAGCCCCG
401 CCCCTGGGTG GAGCCTTCCA AGAGGCCCCA CCCCTCTCGG CTGACCCGGC
451 TCTACCCTGG CCGCAGGAGG ACGGCCGCTT CGCGGGAAAG GAGCTGGATC
501 CCCAGGCAGA TAACTATGTG CCAAACCTGC TGGGCTCTCGT GGAGGAAAAG
551 CTGCTGAAAC TGCAGGCGCA GCTCCAGGGC CACGACGTGC AGGAGATGCT
601 GTGCCACATC GCTAACCGCG AGGTGCCCTG CAGCCATGGG GCAGAGCTCC
651 AGAGATCTAG GAGTAGGGCT GGGTCGGAAA ACCAGGGAAC CGGGGGCAAC
701 TTCGGGAGAC CAGATTGCAG GCTGATCAGC CAGGAAGGGG TCTCGGGGAC
751 AACGGGCGAG GTTTGACGAG CCAGGGGTGA GCCCGAGCCA CGAGGGAAGA
801 AACCAGAAGA GGTGCAGGCG GGTCTGTGA GGCTAGAGGA CGCGGCTGGG
851 GTATCCGGAT GGGGAGAGGC CTAGTGGCAG GGCCCCAGAG ACCAGGGGCG
901 GACCTGAGTG ATGAGGGGGT GAAGCTGGGT ATCAAGGTGC CTGCTTAAGG
951 AGGGATGGGA GTGGGCAGTG GTCTCCAGTT GGAAGTAGGG CATATCTCGG
1001 AATGGGGTAG GGAGGAGGGG GAACGTGACT TCTCGGTGGC CCCTGGCCCC
1051 GCCACCTTC CTCTCTCGCT TCCCATCCCC GCTTAGTTCC TCGCCAGCTT
1101 AGAGGGAAGG CTGCCC GAAT ACAACACCCG CATCGCCCTG CCCCTTGCCA
1151 CTTCCAAGGA CAAGTTTTTT GTTCAGTATG GGTGGGGGGA TGAAGGGGTG
1201 CGGTTGGGGC CCAAGCTGG GTAGCAGGAG GTTGTGAGAT ACCCTGTTCA
1251 GGGCCAAGGT GGGGGTGGAG TGGGGTGGTA AAGGGAAGGG GAGCCAGGGA
1301 TGGGGACCTG GATCCACAT CGCTCCCTGC TCATCCCCC ACCCCCTTA
1351 GACGAAGAGA GTGAGGAGGA GGACAACGAG GTAGTGACCC GCGCATCACT
1401 CAAGATCCGT TCCAGAAAT TAATCGAAAG TCACAAGAAG CACCGTCGCT
1451 CTCGGAGGTC CCGACTCGT CCTGACACCC ACCAGGCGGC CCCTTCGGAG
1501 CCCCCGAATC TCCGGGTCTA GCGCACGCCA CGGGCGCTTC AGGGGCTGAA
1551 CGCGGCCGGA CCGGGAACGG AGGCGGCCAG CGGGCGCCGG AGGGGAGGAA
1601 GGGGCCGGGC CAGACGTTCC CACAGTAAAT CTCCCCAGCT GGGTCCGCCC
1651 CGGCCTCAGA GTTGCGCAAT AAATGTTACC GACCATGCCC CTGGGTATTG
1701 ATCTGTTTTT GACCCTGCAC CACCAAGAG ACGGCTGTCC CTGAAAACCC
1751 AGGGCCACAG ACTGCCTCCT CCAACCTGGG TCATGATGAC TCCCATCAGC
1801 TAGTGACGCA GATGGAGCTT AAAAATGGGA GATGGCCCGA TGAGTGGTT
1851 TTATGCCTGT AATCCCAGTA TTTTGGGAGG CTGAGTTGGG AGGATCACTT
1901 GAGTCCAGGA GCTCCAGGCT GCAGTGAGCT ATGATCGTGC TACTGCACTC
1951 CAGCTGGGC CACAGAGCCA GACCCTGTCT CAATAAATAA AATAAGGGCG
2001 GGGTGCAAGT GCTCATT CAT ACCTATATT CCAGCACTTT GGGAGGCTGA
2051 GCTGGGTGTG TCGCTTGAGC CCAGGGGTTT CAGACTAGCC TGGGCAACAT
2101 GGTGAAAACC AGTTTTTACC AAAAAAAAAA AAAAAAAAAA AAGCTGAGCA
2151 TGGTGGCATA TGCCTGTAGT CCCAGCTACT TGGGAGACTG AGGCAGGAGA
2201 ATGGATTGAA CCCAGGAGGC GGAGATTGCA GTGAGCCAAG ATCAAGCCAC
2251 TGCACTGCAG CTTGGCAAC AGGAGTGAGA CCCTGTCTCT AAAAAATAAT
2301 AAGGCTGGGC GCCGTGGTTC ATGCCTGTAA TCCCAGCATT TTGGGAGGCT
2351 GAGGTGGGCG AATCTCTTGA GGCCAGGAGT TTGAGACCAG CCTGGCAAGT
2401 ATGGCAAAAC CCCGCCTCTA CAAAAAATAC AAAAATTAGC TGAGCATGGT
2451 GGCGGACCTT GTAATCCTAG CTAATTGGGA GGCTGAGGCA CAAGAATCCT
2501 TTGAATCTGG TTAGGAGGAG TTGAGTGAG TTAAGATCAA GCCACTGTAC
2551 TCCAGCATGG GTGACAGAAC GAGACTCCAT CTCAAATAA TAGCAATAAT
2601 AATAAAAAGT GGAAGATGCC CCCCACTTG ATCAAGCTAG CCCCTTCCAC
2651 TGGAGGACAG AGGACTCTGG TCTGGGGACA CACACATGCC CCCACACAGG
2701 AGCTCCCCCA CATCTGGGGA TACAAAAAAG ACCCTTGGG GACAGATATG
2751 TCCTTTCTTC TGGGACAGA TTGATAGGCA CCCAGCGGAA GAGCCAGGAC
2801 CTCTCCTGGG CTGGCGCTGG GTCCGCTGG AGGCACCCAG AGGCTGGGTC
2851 CGGCCTGCCC TGCCCCGCC CGCCCCAGCA GCTCGGCCG TCCGCCCTC
2901 TGGCCTCAGC GCCCGGCCAC TGCCCGCCG CCGCACCCG CCACCCGCCG

```

FIGURE 3A

2951	GCCCTTCCGC	CTCACTCAGC	GGCGCCACTG	AGAGGGACGG	GCGCCGGCCA
3001	TGGAGCGCAC	AGCAGGCAAA	GAGCTGGCCC	TGGTAAGGGG	ACAAGGGATC
3051	CCCGGACCCC	GCATCCCTGG	TGACCCGCAG	GTCCAGAAAC	TCCAAGCGCC
3101	CGCCCGTCGG	ACGGTATCTG	CTCCCAATCT	GAACCTGCCC	TGGAGTCCCC
3151	TCCTGGGGAC	TGCGGGCCCT	TGACCCAGTG	AAGCGACTGG	TTCTCTTAG
3201	GGATGGGGGC	GCGAGTCTCT	GAGCGCAGTC	GGCAGAAAGA	GCTAGAGACA
3251	GGTTCTATTA	GACTGGGCCC	TGGGACATCC	CCAAATGCCA	CCCCATGTCC
3301	TCAGGACCTG	GGAGGAGGGG	ACCCGCAGCG	AGGAGGGGAC	TAGCCTGGGA
3351	CCCCAGCCCT	AGTCTCGCAG	CTTCTGGCCG	GGAAGGGGCG	TGGGGATGCA
3401	GCAGGAGGAC	TCGGCCCGAG	TCCGAGCGGC	CAAGGAGGCT	GAGGCCCCAG
3451	GACCTGTGCC	CCTTTGGTGC	CCTGAGTCCG	CCTGTGCGTC	CAGGCACCGC
3501	TGCAGGACTG	GGGTGAAGAG	ACCGAGGACG	GCGCGGTGTA	CAGTGTCTCC
3551	CTGCGGCGGC	AGCGCAGTCA	GCGCAGGAGC	CCGGCGGAGG	GCCCCGGGGG
3601	CAGCCAGGTG	AGGAGGGGGT	TTGGTGGGTG	GCGCGGGGCC	GGAAGCGACC
3651	AGTTGAGGGC	GGAGCTGGAG	AGCCGAGCAC	AGGCCGCCAG	GTGCAGTGGG
3701	CGGAAGGAAG	GGAGGGGCTC	GGAGGCGACC	AGATGAGGCG	ACCAGGTAGA
3751	AAGGGGACTG	GGGGCGGCCA	GGTAAGTGGG	GGGAGATCCA	GGGAATGGGG
3801	TGGGGCCAGG	CGATGGCCGC	GCA GTTCCCG	AGAGGAGCCT	AGGGACAAC
3851	TGGTAAGGAC	AGA ACTGGAC	GGCAGAGTTG	GGAAAGGCAG	GTTTAGAGGG
3901	CCGGGGCTGG	AAGGTGGAAT	GGGGTTGGTT	TAGCAAGTGG	CTAGGTGAGG
3951	GCGGATGGGG	CAGCCAGTGA	AGCGCGACAG	GAGGGCTGAG	GGAAGCCCTG
4001	GGTGGAAGAA	AGTGTGTGGG	GCGGGGGCGG	GGGGGTGGGG	GGAGGGGACG
4051	GGAGGGGGAG	GGGACGGGAG	AGGGAGTAGG	GGACAGGGCA	TGGGAGAGGG
4101	AGGGTTTCCA	GGGCAAGTTG	CAGGAGCTAT	TTGTGGATGG	GGAGGAACAA
4151	TAAC TTCAAG	CGGGCAGGGA	GTGGGGCACA	CACCTATAAT	CCCTGCGCTT
4201	CGAGAGACCA	AGGCAGAAGG	CCAGGAATTG	GAGACCAGCC	TGGACAACAC
4251	AGCAAGATTG	TCTCTAATAA	AAATAAAAAAT	TAAAAAACTA	GCTGTGCGTG
4301	ATGATGCCCC	GCA GTTGGTCC	CAGCTACTCA	GGAGGCTGGG	GCAGAGGGAC
4351	CGCTTGAGTC	TAGGACTTGG	AGGCTGCAGT	GAGCTATGAT	TGTGCCACTG
4401	CACCCAGGCC	TGGGCAACAA	AACAAGACCT	GTTTCTAAAA	AAAAACAACC
4451	AAAACAATAA	CTCCAAGAAG	CCGGGAGACA	GAGGAATCAC	ATGAAAGAAT
4501	GGTGCTACAG	GCGGGGCGAG	GTGGCTCACG	CCTGTAATCC	CAGCACTTTG
4551	GGAGGCCGAG	GCAAGTGGAT	CATCAGGTCA	GGAGTTCAAG	ACCAACCTGG
4601	CCAAGACGGT	GAAACACCGT	CTCTACTAAA	AATACGAAAA	AACTAGCTGG
4651	GCTTGGTGGC	GGGTGCCTGT	AATCCCAGAT	ACTTGGGAGG	CTGAGGCAGA
4701	GAATTGCTTG	AACCCAGGAG	GCGGAGGTTG	CAGTGAGCCA	AGATCACGCC
4751	ACTGAACTCC	AGCCTAGGTG	ACAGAGTGAG	ATTCTGTCTC	AAAAAAAAAA
4801	AAAAAAAAGT	GGTGCTAGGG	GCTGGGCACG	GTGGTTACAG	CCTGTAATCC
4851	TAGCCCTTTG	GGAGACTTTG	GGAGGCCAAG	GGGGGCAGAT	TACTTGAGGT
4901	CAGGAGTTCC	AGACCACTCT	GACCAACATG	GTGAAACCCT	ATCTCTACAA
4951	AAATACAAAA	ATTAGCTGGG	CTTGGTGGTG	TGCGCCTGTA	GTTTCAGCTA
5001	CTTGAGGGCT	GAGGAAGGAG	GATTGCTTGA	ACCCAGGAGG	CAGAAAGTTGA
5051	AGTGACCCAA	GATCGTGCCA	CTGCACTCCA	GCCTGGGCAA	CAGAGTGAGA
5101	CTCTGTCTCA	AAAAAAAAAA	AAAAAAAAAA	AAGAGTGGTG	CTAGTGATGA
5151	ATGTGACTAG	AGAAGGGGTG	CTGTGAGGAC	CACTCCTGCT	CTCTCATGGC
5201	CACCTCTCCC	CTCCTGCAGG	CTCCCAGCCC	CATTGCCAAT	ACCTTCCTCC
5251	ACTATCGAAC	CAGCAAGGTG	AGGGTGCTGA	GGGCAGCGCG	CCTGGAGCGG
5301	CTGGTGGGAG	AGTTGGTGTG	TGGAGACCGT	GAGCAGGACC	CCAGCTTCAT
5351	GCCCGCCTTC	CTGGCCACCT	ACCGGACCTT	TGTACCCACT	GCCTGCCTGC
5401	TGGGCTTTCT	GCTGCCACCA	ATGCCACCGC	CCCCACCTCC	CGGGTCAGTA
5451	GCGAACCATA	ACCTCCGTAT	TCTCCACCCT	AGAACCCCAA	CTGGGCACCC
5501	CCCTCCACCT	CCTCAGGTGT	GGAACCTGGA	AACACCTCCC	AGACCCAGAG
5551	CCCTCTTCCT	AAGCCCCCTC	TAGGTTCCCC	CTTCTTCACC	TGCTGGGGGG
5601	CCTCTTCCCA	GGGTAGAGAT	CAAGAAGACA	GCGGTACAAG	ATCTGAGCTT
5651	CAACAAGAAC	CTGAGGTGGG	TCCTTCATCC	AGATAGGGGA	GTGCGGGGAG
5701	GGAAATCCAA	GAGGTCAAAG	GTTAGCAGTC	GGACTGGGGT	TTTGAAAATT
5751	GCAGGTTGGG	TAATAAGAGA	CTGGGAGTCA	GGTGGGGCGT	GGTGGCTCAT
5801	GCCTGTAATC	CCAACACTTT	GGGAGGCCGA	GGCAGGTGGA	TCACTGGAGG
5851	TCAGGAGTTA	GAGACCATCC	TGGCCAATAT	GGCGAAACCC	TGCTCTACT

FIGURE 3B

5901	AAAAATACAA	CAACAACAAA	AAAAGGTAGC	TGGGTGTGGT	GGCGCATGCC
5951	TGTAGTCCCA	GCTACTCGGG	AGGCTGAGGT	TGCAGTGAGT	CAAGATCAGG
6001	CCATTGCACT	GCAGCCTTGG	TGACACAGTA	AGACTCTATC	TCAAAAAAAAA
6051	AAAAAAAAAA	AAGGTACCAG	GAGTCATATT	CTATGTCCCC	CACTCTGGAC
6101	CCAGCTCTGA	GACCCTGCCT	CTCTGGCCAG	GGCTGTGGTG	TCAGTGCTGG
6151	GCTCCTGGCT	GCAGGACCAC	CCTCAGGATT	TCCGAGACCA	CCCTGTCCAT
6201	TCCGACCTGG	GCAGTGTCCG	AACCTTTCTG	GGCTGGGCGG	CCCCAGGGAG
6251	TGCTGAGGCT	CAAAAAGCAG	AGAAGCTTCT	GGAAGATTTT	TTGGAGGAGG
6301	CTGAGCGAGA	GCAGGAAGAG	GAGCCGCCTC	AGGTGTGGAC	AGGTGAGGGG
6351	TTTTCAGATC	CAGTCGTGTT	CTGAGAAGGC	CTTTCCTGTC	TGCTTCTTCC
6401	CACACAGGCT	TTCTCTCCCC	TCTCAGAGCT	ACAAAACCTA	AGCAAGATTT
6451	TAAACTCTAA	GCCTCAATTT	CTTCATCTTT	ACAATGGGGA	TAATAATTCT
6501	TTGTACGCG	GGCGTGGTGG	CTCACGCCTG	TATCCCAGCA	GTTTGGGAGG
6551	CCAAGGATGG	TGTATCACCT	GAGGTCAGGA	GTTTGAGACC	AGTCTGACAA
6601	ACATGGAGAA	ACCCCATCCC	TACTAAAAAT	ACAAAATTAG	CCGGGCGTGG
6651	TGGGGCATGT	CTATAATCCC	AGCTATTCGG	GAGGCTGAGG	CAGGAGAATC
6701	GTTTGACCCC	GGGAGGCGGA	GGTTGCGGTG	AGTCGAGATC	GTGCCATTGC
6751	ACTCTCGCCT	GGACAACCAG	AGCGAAACTC	CGTCTAAAAA	AAAAAACAAA
6801	TTCTTTGCT	GAAGTATTAG	CATGTGTCTA	ATACTTTTCC	CTCCTTGGTG
6851	CCGTTGGGTC	AGGATGCTCT	GTGTTTCTAG	CTACAAACCA	TTGCCTTGAT
6901	ACTTGTCTTT	ATTTTCTTTT	TTTTGAGTCA	GGGTCTTGCT	CCGTTGCTCA
6951	GGCTGGAGTG	CAGTGTCTCC	ATCATGGCTC	AGTGCAGGCT	CAACATCCTG
7001	GACTCAGGTG	ATCCTCCCGC	CTGGGTCTCC	AAAACCTGCTG	GCATTACAGG
7051	CGCGAGCCAC	TATACCTAGC	CTGTAAAAAT	TTTCTTATTT	TTGAATTTCT
7101	TTTTAAATTT	AATTTAATTT	AATTTTATTT	TTTTATCTAT	TTTTTTTTTT
7151	AGACAGAGTC	TCGCACTGTT	ACCCAGGCTG	GAGTGCAGTG	GCACAATCTT
7201	GGCTCACTGC	AACCTCCACC	TCCTGGGCTC	AAGCCATTCT	CCTGTCTCAG
7251	CCTCCTGAGT	AGCTGGGACC	ACAGGCGCAT	GTCAACACGC	CCGGCTAATT
7301	TTTTTGTAAG	GGTGAGGTTG	TGCCATGTTG	CCCAGGCTGG	TCTCAAATC
7351	CTGAACCTAA	GTGATCTGCC	TGCCCTTGCC	TCCCAAAATG	CTGGGATTAC
7401	AGCCATAAGC	CATTGTGCAT	GCGTAGCCTC	CTTACTTGAT	TATTGGCTTT
7451	TGCTCATCTC	ATAGGCTGTG	AGTGCATGAG	AGGAGGACCT	GTTGTTCTTG
7501	CTCCCAGCTC	TGTCCCCAGG	GGCAGGAACA	ACACAGATTA	GTTTGCTGAA
7551	TAATTGCATC	CTGCTTAGGA	AGTATCATCT	TTCACCCATC	TGTATTTGAT
7601	CTGATCCACA	TCACAAAAGC	ATCTCTATCC	CTAATCCCCA	TCGCTTAATC
7651	TCCAGATTAT	AGAGGCCACC	TTCTGTCCA	ATTTACAAAG	TAGCAGCCAC
7701	TTCTCTATCC	CTGGTGACAA	AGTCTCAGTT	ATTTATATAT	ATATAAAGGT
7751	ATATATATAT	ATATATATAT	ATATATATAT	ATATACATGA	AGGTGTATAT
7801	ATATATATAT	ATATGAAGGT	ATATATATAT	ATGTATATAT	ATGAAGGTAT
7851	ATATATATAA	AGGTATATAT	ATATAAAGGT	ATATATCTAA	AGGTATATAT
7901	ATATATAAAG	GTATATATAA	GGGTATATAT	ATATAAAGGT	ATATATATAT
7951	ATATGAAGGT	ATATATATAT	ATGTATATAT	ATGAAGGTAT	ATATATATAA
8001	AGGTATATAT	ATATAAAGGT	ATATATATAA	AGGTATATAT	ATATATAAAG
8051	GTATATATAT	ATAAAGGTAT	ATATATATAT	AAAGGTATAT	ATATATAGGT
8101	GTATATATAT	ATATATATAT	ATATATATAT	ATATATATAT	ATATGATTTT
8151	TCCAGCTGAT	TCCAAGTCAT	TAGAGCTCCA	TAGTTCACCTG	TGGTATCCAC
8201	TAGCACCTGT	CGCTATTTAA	ATTAATTAAA	ATTGGCTGGG	CGCGGCGGCT
8251	CATGCCTATA	ATCCCAGCAC	TTTGGGAGGC	CGACGGGGGC	GGATCCCAAG
8301	GTTTCGGAGAT	CGAGACCATC	CTGGCTAACA	TGGTGAAACC	CCGTCTCTAC
8351	TAAAAATACA	AAAAAATATT	AGCCGGGCGT	GGTGGCGAGC	GCCTGTAGTC
8401	CCAGCTACTC	GGGAGGCTGA	GGCAGGAGAA	TGGCGCGAAC	CTGGGAGGCG
8451	GAGCTTGACG	TGAGCTGAGA	TCAAGCCACT	GAACTCCAAC	CTGGGTGACA
8501	CAGCGAGACT	CTGCCTAAAA	AAAAAAAACC	AAAAAACAAA	ATTATAATAA
8551	TAATTAATTA	ATTAATTTAA	ATTAATAAAA	ATTCAGGTCT	TTTCTTTTAA
8601	GAGATGGGGT	CTTGCCATGT	TGCCCAGGCT	GGTCTCGAAC	TCCTGGGCTT
8651	AAGCAATCCT	CCAGCATCAA	CCTCTCAGAG	TGCTGGGATT	GTAAGTGTGA
8701	GCTACTGTGC	CTGACCCTGC	CTTTTTTTTT	TTTTTTTTTT	TTTTGAGACG
8751	GAGCCTCGCT	CTGTCACCCA	GGCTGGAGTG	CAGTGGCGCC	ATCTCGGCAC
8801	ACTGCAACCT	CCTCCTCCCA	GGTTCAAAAG	ATTCTCCTGC	CTTAGCCTCC

FIGURE 3C

8851	CAAGTAGCTG	GGATTATAGG	CACCTGCCAC	CACACCCAGC	TAATTTTGTG
8901	TTTTTATTAG	AGACAGCGTT	TCATATGTT	GGTCAGTCTG	GTCTTGAAC
8951	CCTGACCTCA	GGTGATCCAC	CCACCTCGGT	CTCCCAAAGT	GCTGGGATTA
9001	CAGGCGTGAG	CTACCATGCC	TGGCCCGCTT	TTTTTTTTTT	TTTTTTTCTT
9051	TTTCAAAATC	CAGTCAAGCA	AAGGCAAAAA	TTCAGGTCTT	CAATCCCACT
9101	ACCCACATTT	TGAGTGCTCA	GCCACCACAC	TGGACATAGC	AGATAGATAA
9151	TTTTTCCACC	ATTGCAGAGA	ATTATATGGA	AAGTGCTGCC	CTAGTTTCTT
9201	TGAGGTCAGA	GGAGAAAATT	AACATTTGTT	TAAGACCTTC	TATGTGCTAG
9251	GCCCTGGGAC	ACACTTTATT	TCATTTTATT	TTATTTTATT	TATTTTFACT
9301	TTTATTTTAT	TTTGAGACAG	AGTCTCGCTC	TGTCGCCTAG	GCTGGAGTGC
9351	AATGGCGCGA	TCTTGGCTTA	CTGCAACCTC	CACCTCCTGG	GTTCAAGTGA
9401	TTCTCCTGCC	TCAGCCTCCT	GAGTAGCTGG	TACTACAGGC	GCCCGCCACC
9451	AGGCCCAGCT	AATTTTTTGT	AGTTTTAGTA	GAGACGGGGT	TTCACCGTGC
9501	TAGCCAAGAT	GGTCTCGATC	TCCTGATCTC	GTGATCCGCT	TGCCTCGGTC
9551	TCCCAAAGTG	CTGGGATTAC	AGGCGTGAGC	CACCGCACCC	GACTIONGAAT
9601	TTTATTTTAT	GATACAGGGT	CTTGCTCTGT	TGCCCAGGCT	GGACTCGAAC
9651	TCCTGGGCTC	AAGTGAGCCT	CCTACCTCAG	CCTCCTGAGT	AGCTAAGACT
9701	ACACTTGCAC	CATGTAGTTT	AGAAGAAAGT	AGATGACCAC	CATGCTCATC
9751	TATTTTATTT	TAACAACTTT	ATTTTGGGTT	CACTTTTTGC	TATGGAAAAT
9801	TTCAGACATA	TACAAAAGTA	GAGAGAATAG	TATGAAGAAC	ATTCAGACAT
9851	CCATCACCTA	TCATCAACGA	TGATCAATTT	CACAAAAAAA	TATTTTCAGG
9901	ATGATTTTAA	AACAAATCCC	GGGCTTATGT	CAATTCATAC	ATAAATGTTT
9951	TGGGTACACA	TGTCTGACAA	CAGGCTTACT	TTTTTTTTTT	TTTTTTTTTGA
10001	GACGGAGTTT	CGCTCTTGTT	GCCCAGGCTG	GAGTGCAATG	GCAGGATCTC
10051	AGCTCACCTC	AACCTCTGCC	TCCTGGGTTT	AAGTGATTCT	CCTGCCTCAG
10101	CCTCCCGAGT	AGCTGGGATT	ACAGGCGTGC	ACCACCACAC	CCGGCTAATT
10151	TTCTATTTT	AGTAGAGAGG	GGGGTTTCTC	CATGTTTGTC	AGGCTGGTCT
10201	CGAACTCCTG	ACCTCAGGTG	ATCCGCCAC	CTTGGCCTCC	CAAAGTGTTG
10251	GTATTACAGG	CGTGAGCCAT	GGCGCCCGGC	CCTTTTATTT	TTATTTTTTTA
10301	ATAACCTTCA	TGTTCACTACT	TAAAAAATA	TCAGAAATAT	TTGATATAAA
10351	AAAAATCCAA	TCCAGGCCAG	ATGCAGAGGC	TCCTGCTGGC	GATCCCAGCA
10401	TTTTGGGAGG	CCAAGGCAGG	TGGATGGGCT	TTGAGCCAG	GAGATTGAGA
10451	CTAGCCTGGG	CAACATGTTG	AAACTTTGTG	TCTACAAATA	ATTAGCTGGG
10501	CATGGTGGTG	ACTGCCTATA	GTCCCAGCTG	CTTGGGAGGC	TGAGGCAAGA
10551	GGATCATTTT	AGCCTGGGAT	GGTCAAGGCT	GCAGTGAGCC	GTGATTATGC
10601	CACGTGACTC	CAGCCTAGGT	GACAGAGCGA	GACCCTGCCT	CAAAAAACAGA
10651	AAAAATACCC	AGTCTATATT	CAAAATATTCA	AATCCCCTGT	TTGTGCCTGA
10701	ACCTTTTTTT	GGACACTGGG	TTTTCTATT	TTGCCTGGGC	TGGGCTTGAA
10751	CTCCTGACCC	TCCCACCTCA	GCCTCCTGAG	TAGCTGGGAC	CACAGGTGCC
10801	CACCATGGCA	CCCAGCCCTA	AATTTTCTTT	TGACAGTTGT	TTCTGGCCAG
10851	GTGTTGTGGC	ACATGCCTAT	AGTCCCAGCT	ACTTAGGATG	CTGAGATGGG
10901	AGGATCTCTT	GACTCCGGGA	AATCAAAAAGC	TGCCGTGAGC	TGTGAGCATG
10951	CCCCTGCACT	CCAGGCGATA	GAGCTGGGGG	AAGGAGGAAT	AGTTGTTTCT
11001	TCAAATTGAA	ATCCAAAGAT	CTACTCAAGG	TATTTGGTTG	TTTGCTTCTC
11051	TTTTTTTTTT	TTTTTTTTTT	TTTGAGATGG	AGTCTCACTC	TGTTGTCCAG
11101	GCTGGAGGGT	AGTGGCGTGA	TCTTGGCTCA	CTGCAACCTC	CGCCTCCTGG
11151	GTTCAAGCGA	TTCTCCTGGC	TCAGCCTCCT	GAGTAGCTGA	GTTTACAGGT
11201	GCCCACCAAC	ACGCCAGCT	AATTTTGTGA	TTTTTAGTAG	TGAGGGGGTT
11251	TCACCATGTT	GGCCAGGCTG	GTCTTGAAC	CCTAACCTTT	AGTGATCTGC
11301	CCACATCGGC	CTCCCGAAGT	GTCCGGATTA	CAGACATGAG	TCACCACGCC
11351	CTACCGGTG	TTTGTTTATA	AGTCTCTTTT	ATTCTGTAAC	AGATCCCCCT
11401	TGCCTCTTGT	TTGAAGCCAT	TAGAGGGCAA	AAAAAATGGG	TCATTTTTCC
11451	TGAGGTATGT	CTCACATTCT	TTTCGACTTA	CCTCATGGTT	TCATGCAGCA
11501	TGTTTCTCTA	TCCCCATAAT	TGCTGTAAGA	TTTAAAGGTT	TGATTAGATG
11551	TAGGGCATTT	TTTTTTTCCAG	GGCCCACTTT	TTTTTGGGGT	GGGGGGAGGA
11601	GAGACAGTTT	CTTGCTCTGT	CACCCAGGCT	GGAGTGCTAT	GGCATGATCA
11651	CAGCTCACTG	CAGCCTTGAC	CTCCTGGGCT	CAAGAGATCC	TCCCTCCTAA
11701	GCCTCTTGAG	TAGGTGGGAC	AGCAGGTGTG	CATCAGGATG	CGCAACTTTA
11751	AAAATTTTTT	TATGTAGACA	TGGGTCTCA	CTACGCCGCC	CAGGCTGGTC

FIGURE 3D

11801	TCAAACCTCT	GGTCTCAAGC	AATCCTCCTA	CCTCAACCTC	CAAAAGTGCT
11851	GGGACTATAG	GTGTGCCCAG	CCCAGTACCC	ACTTCTAAAA	ACTAATATTT
11901	TGCAATGCCA	CCTGTCCTAA	TTCAAGATGA	AAGAGGTAAT	TACACAGATT
11951	TACAAAGATT	ATTTTAAAA	AATAGTATTG	GGGCAGGGTG	CTATGGCTCA
12001	TGCCTGTAAT	CCCAGCACGT	TGGGAAGCCG	AGGCAGGAGG	ATCACCTGAG
12051	GTCAAGAGTT	CGAGACCACT	CCGGCCAACA	TGGTGAAACC	CCATCTCTAC
12101	TAAAAAAAAT	AAAAAATAAA	ATAAAAATAA	ATAAAAAATA	AATAAATAAT
12151	AAAAAAATAT	ATATATATTT	AAATTAGCTG	GCTGGGCATA	GTGGCACCTC
12201	CTGTAGTCCC	AGTTGCTCAG	GAGGCTGAGG	CAGGAGAATT	GCTTGAACCC
12251	TGGAGGCAGA	GGTTGCAGTG	AGCCGAGATC	GAGCCACTGC	ACTCCAGCCT
12301	GGGCGACAGA	GCAAGACTCC	ATCACAAAAT	AAAAAAATAA	AATAAAATAA
12351	TAGTATGATG	CCATAACTAG	TACAAAGGAG	AAGGAAAGTG	AGAGTAACTT
12401	ACACAGCAAT	AAACCATGTT	TTCAATGGGT	AATGCTTGGG	TATGCCCCAC
12451	TAGGACACAT	GATGAGGTTG	TCCCGTGTCT	TTGCCTGTCC	TAGCGTCACA
12501	GTAGAGTGTC	ACGGTGCTGT	TGTAAGTACA	GCAACAAGCA	CCAACGAACG
12551	CACAGGAGGG	CACTGGTGAG	GCAAAGACAG	CAACATAGGT	TCTGGGGACA
12601	TCATTTTCCA	CACTTGTGAA	CAACATTTGC	AATTTGCAAA	CAAAACAAAG
12651	CCCAGACTTT	CGTGGTCTTT	GCATTCTTGG	AGCCAAAAAA	ATTTGTGTTT
12701	ATGAACAAAA	TAGTCAGGTT	CTAGGTGCAT	ATTATTGCAA	ACATGTTTTT
12751	CTTTTCTTTT	TGTTTTTGT	TGTTTGTTTG	TTTTGTTTTG	TTTTGTTTTT
12801	TGAGATGGAG	TCTCGCTCTG	TCGCCCAGGC	TGGAGTGACG	TGGCATGATC
12851	TCGGCTTACT	GCAAGCTCCG	CCTCGCCGGT	TCACGCCATT	CTCCTGCCTC
12901	AGCCTCCTGG	GTAGCTGGGA	CTACAGGCGC	CCGCCACCAC	GCCTGGCTAA
12951	TTTTTTCTAT	TTTTTAGTAG	AGACGGGGTT	TCACCATGTT	AGCCAGGATG
13001	GTCTCGATCT	CCTGACCTCG	TGATCTACCC	GCCTTGGCCT	CCCAAAGTGC
13051	TGGGATTACA	GGCGTGAGCC	ACTGCCCCCG	GCCTTCTTTT	CTTTTCTTTT
13101	TTTTTTTTTT	TTGAGACAAA	GTCTCTGTCA	CCCAGGCTAG	AGTGCCGTGG
13151	CGTGGACCTG	GCTCACTGCA	ACCTCCACCT	TCTAGGTTGA	GGTGATTCTC
13201	TAGCCTTAGC	TCTCCGAGCT	GGGATTACAG	GCACTTGCCA	CCATGCTCAG
13251	CTGATTTTTG	TATTTTTAGT	AGAGACAGGG	TTTCGCCATG	TTGGCCCGAC
13301	TGGTCTCGAA	CTCTTGACCT	CAAGTGATTG	GCCTGCCTTG	GCCTCCCAAA
13351	GTGCTGGGAT	TACATGTGTG	AGCCACTGTG	CCAGACCCCT	TCTTCCTTTC
13401	TTAAAGACAA	GTCAAGTGCA	GTAAGTGAAG	GGGGGGAAAG	AGTAGAACAA
13451	GGAGTTCGAT	CTGTAAGTGT	GAACAATCAA	TTGAGATAAG	TCACTACCTT
13501	GGGACCAAGC	ACAAACAGGT	TTTTCAAAGA	CACAAATGTC	TGGAGATACA
13551	TTTGAGGGCT	AGAGGGCACA	ATTCAAGATC	CCAGTTTCCA	AAGTTTCCCC
13601	TCCAGGGTGC	CACCATCAAA	ATCCACTAAA	GTAAATTTAT	TCATATTTGT
13651	TCAGCACTTT	ATAGCAGTCT	GGTAGCATGA	TCTTTTTTTT	TTTTTTTGAG
13701	ATGGAGTCTC	GCTCTGTGCG	CAGGCTGGAG	TGCAGTGACA	CGATGTGGGC
13751	TCACTGCAAG	TCTCGTCTCC	AGGGTTCAAG	CGATTCTCCT	GCCTCAGCCC
13801	CCCGAGTAGC	TGGGATTACA	GGCGCGTGCC	ATCACGCCCG	GCTAATTTTT
13851	GTATTTTTTT	TTAGTAGAGA	CGGGGTTTCA	CCGTGTTGGC	CACGCTGGTC
13901	TCGAACCTCT	GACCTCAGGT	GATCCACCCG	CCTCGGCCTC	CCAAAGTGCT
13951	GGGATTACAG	GCGTGAGCCA	CAGCGCCCGG	CAGCATGATC	TTAAACGAAA
14001	ACAAAAACGA	AATCCACAGC	CAGGCGCACT	GGCTCACACC	CGCAATCCCA
14051	AAACTTTGGG	AAGCCAAGAG	GGAGGATCGC	TTGAGCCCAG	GTGTTTGAGA
14101	CCAGCCTGGC	AACATAATGA	GACCTGTCT	CTACAAAAAA	TAAAAAATTA
14151	GCTGGGCATG	GTGGTGTGTG	CCTATAGTCC	CAGCAACTCA	GGAGGCTGAG
14201	GCAGGAGGAT	CACTGGAGCC	CAAAAGGTTG	AGGCTGCAGT	GAACTGTGAT
14251	CACACCACTG	TACTCCAGCC	TGGGTGACCA	AGGGAGAGCC	TGTTTCAAAA
14301	AGAAGGCACA	GCTTACCCCT	GCAATCCCAG	CACTTTGCGA	AGTCGAGGCA
14351	GGCAGATCCC	TTGAGGTCAG	GAGTTCAAGA	CCAGCCTGGC	CGACATGGTG
14401	AAACCCTGTC	TCTACAAAAA	TACAAAAGTT	AGCTGGGCGT	GGTGGCTCAG
14451	TGCCTGTAAT	CCCAGCTACT	TGGGAGACTG	AGGCAGGAGA	ATTGGTTGAA
14501	ACCTGGAGGC	GGAGGTTGCA	GTGAGCCAAG	ATCACGCCAT	TGCATTCCAG
14551	CCTGGGCGAC	AGAGTCAGAC	TCCGTCTTAA	AAAAAAAAAA	AAAAAGGCAC
14601	AGAGAGGTTA	AAATACATGC	TCTACACAGC	AAGCTAGTGG	ACGAGTTTGC
14651	ATCTGAGTTT	GAGACTTTCT	GACAATAGCC	TTCCCTGAAC	CAGGAAGTGC
14701	TATCACCTCT	TTCCAAAAAA	AAGAGGTCAG	ATTAATCTTA	TCCTAATACA

FIGURE 3E

14751	TGTTAAAAAT	CATAAAGCTC	TATTTTCTTC	TCTGGCCTTT	GAGTACCCGG
14801	CTTCAAACCC	CTGCCCTGCC	ATTTACCAAA	GGTGTGACAA	ATTGTTCTTT
14851	GCCTCCCTTT	CCTTAATTGT	AAAAGGTGGA	TAAATAATAG	TACCTCCCTC
14901	ACTGGACTCA	CAGTAACTCA	GTGGTGAGTT	ACTGAGTAAA	TCCACACTAG
14951	CTGCTTAGTG	AACATTACTG	TTGCTGTTAC	ATCCTTAAAA	ACACTCAGGG
15001	CCAGGCGTGG	TGGCTCACAC	CTGTAATGCC	AGCACTTTGG	GAGGCCAAGG
15051	CGGGCAGATC	ACTTGAGGTC	AGGAGTTTGA	GACCAGCCTG	GCCAACATGG
15101	TGAAAGCCCG	TCTCTATTAA	AAATACAAAA	ATTAGCCGGG	CATGGTGCCA
15151	CATGCCTGTA	ATCCCAGCTA	CTCAGAAGGC	TGAGGCAGGA	GAATCACTTG
15201	AACCCAGGAG	GCGGAGGTTG	TGGTCAGCTG	AGATTGCGCC	ATTGTA CTCC
15251	AGCTTGCGCA	ACAGAGTAAG	ACTGTCTCAA	AAAAAAAAAA	AAATTTAAGA
15301	GAGCTCTCCG	TTTTACAAAT	GAGGAAAGTG	AGCCTCAGAG	AGGGACAGGG
15351	ACTCACCCAA	GGTCACACAG	CCAGTCTTGG	ATTCAAACTT	GAGAGTTTGT
15401	AACCTTTTCT	AATGATCAGG	ACCTCCCAGA	GTTGCCCAAA	CTTCTGACCC
15451	AGACTCTTCA	GAGGCCTGCG	CGGAGGAAGA	GGAAGGGCTC	ATGCCTCAAG
15501	GTCCCCAGCT	CCTGGACTTC	AGCGTGAGCG	AGGTGGCCGA	GCAGCTGACC
15551	CTCATAGACT	TGGTGAGGAT	CCCGGACAGG	GTGCGGATGA	GCCACAGTGA
15601	GGGGACAGGT	TCTGCTAAGC	ACCAATCCCA	CACCCCTCCC	CTGGCCCAGG
15651	AGCTCTTCTC	CAAGGTGAGG	CTCTACGAGT	GCTTGGGCTC	CGTGTGGTCG
15701	CAGAGGGACC	GGCCGGGGGC	TGCAGGCGCC	TCCCCCACTG	TGCGCGCCAC
15751	CGTGGCCCAG	TTCAACACCG	TGACCGGCTG	TGTGCTGGGT	TCCGTGCTCG
15801	GAGCACCGGG	CTTGCCCGCC	CCGACAGAGG	CGCAGCGGCT	GGAGAAGTGG
15851	ATCCGCATCG	CCCAGGTGTG	TTGCGGGCGC	GGAGAGGGGA	TGCGGGGGCG
15901	GGCCCTGGGG	CAAGGGGAAA	AAATGAGGGC	TCCGGAGAGA	GATAGGGGCG
15951	AGTCTAGCGG	AGGGAGGGAA	CGGGGTGGAA	AGTTGATACC	TAGGGTGAGA
16001	CTTGGGTTCA	GGGAGGAGGG	TCTGGGTCCT	GCAGAGAGGC	CGCGGGCACG
16051	ACTAGTCCC	AAGGGAGCTG	GGAGAAGTAG	GGAGCCCAGA	CCGGAGAAGT
16101	CAAGTCCGA	GGCAGGGGCT	GGAGGGGACG	CTGGGGAGGG	GCTGGAGCCC
16151	GAGGGAGGAG	GGAGGAAGGG	AATCCTAGGG	AATAAGTGGG	AGTCTTGTA
16201	GCTTGTCGGA	TGTGAGACAA	CACCCAGGGG	TCCGACCTGG	CGTCACAAGT
16251	CACGGGATCA	GGCTGGGCGC	AGTGACTCAC	GCCTGTAATC	CCAGCACTTT
16301	GGGGAGAGGG	AGGATCGCTT	GAGCCCTTGA	GTTTGAGACC	AGCCTAGGCA
16351	ACATAGTGAG	ACCAATGTTT	CTAGAAAAAA	AAAAAAAAATT	AAAAAAATTA
16401	AAAATGAGAC	TTACAAAAAA	ATTAGCCGGG	TGTGGTGGTG	TGCCCCTGTA
16451	ATTCCCAGCT	ACTTGGGAGG	CTGAGGCAGG	ATAATCACTT	GAACCCGGGA
16501	GGCGGAGGTT	GCAGTGACTC	GAGATCGGGC	CACTGCATTG	CAGCCTGGGG
16551	GATAAAGCGA	GACTCTGTCC	AAATAATTAA	TAATAATAAT	AATAATAAGC
16601	CATGCATGAT	GGCGCGCGCC	TGTAGTCCCA	GCTACTGAGG	CAGGAGGCTG
16651	AGGCATGTGG	ATCGCTTAAG	CCCAGGAATT	CCAGGCAGCA	AGTGAGCTAT
16701	GATCGAGCCC	GTCGACTCCA	GCCTGGGCCA	CAGACCCTAT	TTTTCAAAAA
16751	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAGAT	GAAGAAGTTT
16801	CAGGATGAAA	GGTGGATAAT	GCCTGGGTCT	GACCTGCGTC	CCCACCGCCT
16851	GGCAGCGCTG	CCGAGAACTG	CGGAACTTCT	CCTCCTTGCG	CGCCATCCTG
16901	TCCGCCCTGC	AATCTAACCC	CATCTACCGG	CTCAAGCGCA	GCTGGGGGGC
16951	AGTGAGCCGG	TGAGCTGGGG	CGGGACCTGT	TCCCCAGCCC	ATCCCAGGTC
17001	TGACCCTCCC	AAGCCACTGA	CCCCTGACCA	CCCTTCTCCT	GTCCTTCCAG
17051	GGAACCGCTA	TCTACTTTCA	GGAACTTTC	GCAGATTTTC	TCCGATGAGA
17101	ACAACCACCT	CAGCAGCAGA	GAGATTCTTT	TCCAGGTAGA	GATGGATGCA
17151	GACTCCAGGG	ATTTTAGGCC	CGGGAAGTCG	GGGGAGGGAC	TTGGGGCCAG
17201	GCAGGGGTAA	TCTCCCTGCT	ATAGTCAGGA	CACTCTGTCC	TTCCCTACCG
17251	CTCAGCAAT	ACCTTATCCT	TGTCCCTGGC	GGGTGACAG	TTTTTCTTTC
17301	CTCTACTTCC	TTGCTTATAG	TTGACTGTCA	GTGACTGCC	TATTTATTCA
17351	CTCAGCAAAA	CACAAGAAGT	CACAAAGAAA	AGGTACTTAA	AGGCCAGAGT
17401	CATAGCACAG	GGTGGGAACA	AAAAAATGT	TCTGAGGACT	TTACCTTGAT
17451	AAGCAAAACT	AAAAAATGTG	TGTCAAAAGT	CTGGCTTATT	TATAGGCAAG
17501	ATTTAGATTG	TCATTGCAAT	CAGGCGCTGG	TTTTTAGAGT	GAATCTAGAA
17551	TGGATCCCTG	GGCCTGGAAC	ATTCTCCACC	CCTCCAGGTT	TGCATGCAAC
17601	TTGCTCACTC	ACCTCCTTCT	GGTCTCTGAT	TAAATGTCCC	TGCCTCTGAG
17651	AGGCCTTCCC	AGCCTCCATC	ATCCCCAAAA	CCACACATCT	GGTTTTTTGT

FIGURE 3F

17701	TGTTGTTGTT	GTTGTCGTCA	TCATTTGTTT	TTTTGTTTCT	TTGTTTGTTT
17751	GTATTGAGAC	AGAGTCTCGC	TCTGTCACCC	AGACTGGAGT	GCAGTGGCAC
17801	GATCTTGGCT	CACTGCAACC	TCCACCTCCC	AGGATCAAGC	AATTCTCTCT
17851	GCCTCAGCCT	CCCATGTAGC	TGGGATTACA	GGCACCACCC	ACGACGCTTG
17901	GCTAATTTTT	GTATTTGGTA	GAGACGGGGT	TTTGCCGTGT	TGGCCAGGCT
17951	GGTCTCGAAT	CCCTGACCTG	AGGTGATCCA	CCTGCCCTTG	CTTCCCAAAG
18001	TGCTGGGATT	ACAGGCGTGA	GCCATCACTC	CCAGCCAAAT	TTACCTGGC
18051	TAACAGAGTG	AAACCCTGTT	TTCCTGCCCA	GCACCTAGAA	CAGCACGTGA
18101	GCTGGGCTCA	GTGACTCAGC	CCTATAATCC	CAACACTTTG	GGAGGCCAAG
18151	GTGAGAGGAT	CACTTGAGCC	CAGGAGTTCA	AGACCAACCT	GGGCAACATG
18201	GCAAAACCCC	ATCTCTGCAA	AAAATACAAC	AATTAGCTGG	GCGTTTGTGG
18251	TGCACGTCTG	TAGTCCCAGC	TATTCAGGAG	GCTGAGGAGG	GAGGATCGCT
18301	TGAACTTGGG	CGGTCAAGCC	TTCAGTGAGC	CAAGATCAGG	CCACTGCACT
18351	CCAGCCTGAG	TGACAAAGTG	AGACTCCATC	TCAAAATAAA	ATGAAATAAA
18401	AAGTAAGTAA	ACAACAGCAA	ATTCAAGGATA	CCCAGGAGAT	CCCTGGCAGG
18451	CCTGTGCCAT	CCAGCTGCGG	ACAAGGATTC	TCTCCTTGTT	AAGGCCAGCC
18501	CTGGGGGCCA	CTACCCACAA	GCCCCACCTC	TCATGGGGCC	TGCTCCCTGC
18551	TGTTTATCTC	CTCCCTACCC	TCATCCAAGG	TGGTCTGGCT	TCTAGAGTGG
18601	GCCTTAACCC	CTGGCTTCTT	TTTTTTTTTT	TTTTTTTTTT	TGAGATGGAG
18651	TTTTGGTCTT	GTTGCCCAAG	CTGGAGTGCA	ATGGTGCGAT	CTTGGCTCAC
18701	TCCAACCTCC	GCCTCCCGGG	TTCAAGCGAT	TCTTCTGCCT	CAGCCTCCCG
18751	AGTAGCTGGG	ATTACAGAAA	TATGCTACCA	TGCCCAGCTA	GTTTTTTATA
18801	TTCTTAGTAG	AAACAGAGTT	TCACTCTGTT	AGCCAGGCTG	GTCTCAAAC
18851	CCTTACCTCA	TGTGATCCAC	CAGCCTCGGC	CTCCCAAGTG	CTGGGATTAC
18901	AGGCGTGAGC	CATCGCACCT	GGCCTACCAC	TGACTTTTGA	TFACTCAAAG
18951	CATGAAGGGT	ATATATGATG	GGTCTGCAGG	CATCGTTCCT	GAGGAATTGT
19001	CCAAGGAGAC	CCCAGACCTG	GCTCAGTTTT	TCTCTTCCCT	CAGGAGGAGG
19051	CCACTGAGGG	ATCCCAAGAA	GAGGACAACA	CCCCAGGCAG	CCTGCCCTCA
19101	GTAGTGTGAT	ACAGTTTGGG	ATGGGGACAA	GTGGGACCTT	CAGGGAGGGT
19151	TGTGGATGGT	GATGGGGTCA	GTAATGGCCC	CAAGTGACTG	GAGCTTTGGG
19201	GGCTGCAGAA	ACCACCCCCA	GGCCTGTGCC	CCTACCTTGG	CACCTTCCTT
19251	ACGGACCTGG	TTATGCTGGA	CACAGCCCTG	CCGGATATGT	TGGAGGTCTG
19301	ACCCCTGACC	CTTGACCCCT	GACCCACAGC	CCACTTGCCC	CCAGCACAAT
19351	GGGCCTCCCA	ATATCCACCC	TTGATCCTAC	CTGTACTCCT	GACACCACCC
19401	CACACTCCCT	TACTACAGTG	GGGCTCCTGA	CATCCCAGCC	CCTGACCTTG
19451	ACCCTTGACC	CTTGACCCTG	GGTGCTGCAA	TTCAGACACA	CTTTGCCCCC
19501	AGGGGGATCT	CATTAACCTT	GAGAAGAGGA	GGAAGGTGAG	TGGAGGCTAC
19551	AGTGGGTGTG	GTGGTGCTTG	AGGGTGGGGG	TGGGGCAGGG	GTAGGGTCTT
19601	AGAGGTCGT	CCTCCAGGAG	TGGGAGATCC	TGGCCCGCAT	CCAGCAGCTG
19651	CAGAGGCGCT	GTCAGAGCTA	CACCCCTGAGC	CCCCACCCGC	CCATCCTGGC
19701	TGCCCTGCAT	GCCCAGAACC	AGCTCACCGA	GGAGCAGAGG	TGACCACCCT
19751	GTAGCCTGTC	CCAGCCCCAC	CCCAGCTGAG	CCTGGGTGAC	CAACTGGATT
19801	CCACCCACTC	CATACACACC	TCCAGCTCCT	CCCAAGACCC	CCTCTTGAGC
19851	CCTGATCCCC	CACTACAACC	TGTGACCTTG	CAGTATCTCC	AGTCGAATCA
19901	AATAGACTGG	GCCTGGTGGT	TACTCGTGT	AATCCCAGCA	CTTGGGAGGC
19951	CAAGGTGGGT	GGATCACTTG	AGCCCAGGAT	TTCGAGACCA	GCCTGGGCAA
20001	CATGGCGAAA	CCCCATATCT	ACAAAAAAAT	ACAAAAATTA	GCTGAACGTG
20051	GCTGGGCACG	GTGGCTCACA	CCTGTAATCC	CAGCACTTTG	GGAGGCCGAG
20101	GCGGGTGAAT	CACATGAGGT	AAGGAGTTTG	AGACCAGCCT	GGCTAACAGA
20151	GTGAAACCCC	GTCCTCTACTA	AAAAATACAAA	AAAAAAATTA	GCCAGGTGTA
20201	GTGGCAGGCG	CCTGTAGTCC	CAGCTACTTG	GGAGGCTGAG	GCAGGAGAAT
20251	GGCGTGAACC	CGGGAGGCAG	AGCTTGACAG	GAGCCGAGAT	GGTGCCACTG
20301	CACTCCAGCC	TGGGCAACAG	AGCGAGACTC	CGTCTCAAAA	AAAAAAAAAA
20351	AAAATTAGCT	GAATGTGGTG	TTGAGTGCCG	TTGGTCCCAA	CTACTTGGGA
20401	GGCTGAGGTG	GGAGGATTGC	TGGAGCCTGG	GAGGCAGAGG	TTGGAGTGAG
20451	CCAAAATCAC	GCCACTGCAG	TTCCAGTCTA	GGTGACAGAG	TGAGACCCTG
20501	TCTCAAAAAA	AAAAAAAAAA	AAATAGTCAC	AATTGACCTC	TGACCTCAAT
20551	TTCAACCCCA	TCTGATTTTC	TGACCTCAAC	TTTAGCATTC	AGCTGGCCAT
20601	TCAACTCAAC	TGTCCCCTCT	GTTGACTTCC	CCATCTTTGG	TCCTATCTGA

FIGURE 3G

20651	CCCATGACCT	TATTCATGAC	CCCTCATCTG	ACTCTCTGAC	CCCAACCCTT
20701	GACCCCTCAGT	TCTGAGTAAC	TGACTCCAAC	TTTTATGTTT	GACTGTCCAG
20751	CTTGACTATG	ACAACGTGTG	CCTTTCCTTC	TATATAACTG	TGACCCTAAC
20801	CATTGACCCC	AATGGTGACC	TGACCCAGT	CTGACCCTGA	CTTTATTTTA
20851	TTTATTTATT	TATTTATTTA	TTTATTTATT	TATTTATTTA	TTTTTGAGAC
20901	AGAGTCTGGC	TCTGTTTCCC	AGGCTGGAGT	GCAGTGGAGT	GATCTCGGCT
20951	CACTGTAGCC	CCCCCCTCCC	AGGTTCAAGC	AATTCTACTG	CCTCAGCCTC
21001	CCCGGTAGCC	GCAATTACAG	GCGCGAGCTA	CCACACCTGG	CTAATTTTTT
21051	TTGTATTTTT	AGTAGAGACG	GGGTTTCACT	ATATTGGCCA	GGCTGGTTTC
21101	GAACCTTTGA	CCCGAAGCAA	TCCTCTCGCC	TCAGCCTCCC	AAAGTGCTAG
21151	GATTACAGGC	GTGAGCCACT	GCACCCAGTC	CTGACCCTGA	CTTTAATCCT
21201	GACCCAATTT	GATTCTTAG	TGCCACCCTG	TGAATCTCTT	TGTGACCTCC
21251	TGACCAGCCA	TCCTGTCCCA	TCTCTGATAA	GACCTTGATG	CTCAATGACC
21301	CTCATTTACC	ACCCTGACCC	TGGCATGTGG	GGTGCCACCT	CTGGCTGCTC
21351	CCCCTTACAC	CCCAAACCCA	CCTCCCAACT	GATTCCAACT	CTTATCTCTC
21401	CATCCCCGTG	ATTTCTCGCC	CCCACCACCT	CATCCACATA	TTGACCCCTC
21451	AGCTACCGGT	TCTCCCGGGT	CATTGAGCCA	CCAGCTGCCT	CCTGCCCCAG
21501	CTCCCCACGC	ATCCGACGGC	GGATCAGCCT	CACCAAGCGT	CTCAGTGCCT
21551	GAGTCTCGGG	GTGTGTGTAG	GGGCGGTGAT	GTGGGCAGAT	ATCAGCAAGG
21601	GCTGCTCCTG	CCTTAGCCTC	ATCCCCGTGC	CCCATCCTTA	GGAAGCTTGC
21651	CCGAGAGAAA	AGCTCATCAC	CTAGTGGGAG	TCCCGGGGAC	CCCTCATCCC
21701	CCACCTCCAG	GTGAGCATTG	TGCTTGGTGA	TGGGACTGGG	GATCATGGGA
21751	TCAGGAGTCA	GCACAGCCAC	CCCACCTCAG	CCTCTGCATC	TCCCCCAGTG
21801	TGTCCCCAGG	GTCACCCCCC	TCAAGTCCTA	GAAGCAGAGA	TGCTCCTGCT
21851	GGCAGTCCCC	CGGCCTCTCC	AGGGCCCCAG	GGCCCCAGCA	CCAAGGTACC
21901	AAGACGGCTT	GTGTGTGCAT	GCGGGCCTGC	GGGCACCCAG	GCTCTGTGTG
21951	TGTGCACGTG	TGTGTGCATG	CACATGTGTA	CACACAGGAT	TGTGGGGCCA
22001	GGAGTGTATA	CAGGAGGCAC	ACTGAGCGCC	CGGGGTATCC	ATCCAGGGGA
22051	TTGCATGCAT	CTGCACGGCC	CTGTTTGGGT	GATCACTCAT	AAATCCGACT
22101	CGTGCTCAGA	TTTGGACCTG	TGTAACCTGCT	TGCCCATGGG	TCATCTAGGG
22151	TGCAATCACA	TCACACCCCT	TTTTATTTGA	AACAGGGTCT	TCTTGCTCTG
22201	TCACCCAGGC	TGAAGTGCAG	CGGTGCAATC	TCAGCTCACC	GCAACTTCCA
22251	CCCCTCCCCC	AGGCTCAAGC	AATCCTTCCA	CCTCAGCCTC	CCAAGTAGCT
22301	AGGACCACAG	GTGTGCACCA	CCATGCCCTG	CTATTTTTTT	TATTTAGTAG
22351	AGATGAGGTT	TCGCCATGTT	GCCCAGGTGG	GTTTCGAACT	CCTGAGCTCA
22401	AACAATGCAC	TCACCTCGGC	CTCCCAAAGT	GCTGGGATTA	CAGGTGTGAG
22451	CCACCGCACC	CAGCCTACAC	TTTTTTGAGG	ACATGTATGT	CCCTAAGAAT
22501	CTGCATACCA	TGGCAGACAC	GGTGGCTATT	GCCTGTGATC	CCGGCACTTT
22551	GGGAAGCCAA	AGTGGGAGGA	TTGCTTGAGG	CCGGGAGTTC	AAGACCAGCC
22601	TGGGCAACAT	AGTGAGCCCT	TATTTCTATT	AAAAGTCAAA	AAAATTAGCT
22651	GGGTGTAGTC	CCAGCTACTC	AGCAGGCTGA	GGTGGAAGGA	TCGCTTGAGT
22701	TTGAGGCTGC	AGTGAGCTAC	GATCATGCCA	CGGCACTCTA	GCCTGCATGA
22751	TAGAGCGAGA	TCCTGTTTAT	GAAGAAAAAG	AGACTGGGCA	CGGTGGCTCA
22801	CGCCTGTAAT	CCCAGCACTC	TGGGAGGCCG	AGGTGGGCCG	ATCACGAGGT
22851	CAAGAGATCG	AGACCATCCT	GGCCAAACGTG	GTGAAACCCT	GTCTCTACTG
22901	AAAATAGAAA	AATTAGCTGG	GTATGGTCGC	GCACACCTGT	AGTCCCAGCT
22951	ACTTGGGAGG	CTGAGGCAGG	AGAATCACTT	GAACCCAGGA	GACGGAGGTT
23001	GCAGTGAGCT	GAGATGGTGC	CACTGCACTC	CATCCAGCCT	GGTGACAGAG
23051	CGAGGCTCCC	TCTAAAAGAA	AAACAAAAAA	AGAAAAGGAA	ATGAAGGAAA
23101	TGAAGGCTGG	GCATGGTAGC	TCATGCCTGT	AATCCCAGCA	CTTTGGGAGG
23151	CCGAGGCCAG	TGGATCATT	GAGGCCAGGA	ATTTGAGACC	AGCCTAGCCA
23201	ACATGGTGAA	ACCCCGTCTC	TACTAAAAAT	AAAAACATTA	GCTGGGCATA
23251	GTGGCACAGG	CCTGTAATCC	CAGCTACTTG	GGAGGGTGAG	GCATGAGAAT
23301	TGCTTGAACC	TGGGAGGCAG	AGGTTGCACT	GAGCTGAGAT	GGCACCCTG
23351	CATTCCAGCC	TGGGTGACAG	AGCAAGACTC	TGTCTCAAAA	AAAAAAGAAA
23401	AGAAAAAGAA	AAGAATCTGT	GTACCAGAAG	AGGAAATGTG	GGCCTGAGTA
23451	TTCATGAGAT	CATGTGTGGG	GTTGTTTATT	GGCATGGGCT	GTGGGTGTAT
23501	AACCGCTGTC	AGCATATGTA	TGTACACAGG	ATTTCTTGTT	TATGAGCATG
23551	GGTTGTGTGT	ATATGGACAC	TGTTTCATGTC	TGTTTCTATA	ACAGGTAACC

FIGURE 3H

23601	AAAGTCTGTA	TATGGTAGGG	TGGTGTATAT	GCAGGCTTGT	GAATGTACTC
23651	CAGTTGCATG	TCCCAGGCTC	TGCATGTGTA	GGGGGTAGTA	GTATGTTTTT
23701	TTGAGATTTT	ATTTTATTTT	ATTATTTATT	TATTTATTTT	TGAGATGGAG
23751	TCTTGCTCTG	TCACGCAGGC	TGGAGTGCAG	TAGCGTGATC	TTGGCTCACT
23801	GCAACCTCTG	CCTCTCAAGT	TCAAGTGATT	CTCCTGCCTC	GGCCTCCCAA
23851	GAGCTGGGA	TTACAGGCAT	GCGCCACCAG	GCCCTGCTAA	TTTTTGTATT
23901	TTTAGTAGAG	ACGGAGTTTC	ACCACGTTGG	CCAGGCTGGT	CTTGAACTCC
23951	CGACCTCAAG	TGATCCGCCC	ACCTCGGCCT	CCCAAAATGC	TGAGATTACA
24001	GGCATGAGCC	ACTGCGCCCA	GCCAATGTTT	TCTTGAGATT	TTAAATGTGG
24051	GGCTATTGAA	TGCACCACTG	GTGGCTGGGG	TGTTCTGTCT	TTTCTAGCCC
24101	TCAGCATCTG	CAGATGGGCC	AAGCTGTAGC	CTCCACCCCT	TACTGCCTGC
24151	AGCTGCCCTT	GAGCCTGGAC	CTGCCAGGCC	CCCGGCCCTT	CGCTTTGCCT
24201	CTGGGCAGCC	CTCGAATCCC	CCTCCCGGCG	CAGCAGAGCT	CGGAGGCCCG
24251	TGTCATCCGC	GTCAGCATCG	ACAAATGACCA	CGGGAACCTG	TATCGAAGCA
24301	TCTTGGTGAG	GGGCTGGGCT	GGGGGTCTGC	TGGAGGCTGC	CCTGCCCTTG
24351	GGGCCGGGGC	CCTCACCTCA	CCTCCCGCCC	CTCTCTTCCA	GCTGACCACT
24401	CAGGACAAAG	CCCCCAGCGT	GGTCCGGCGA	GCCTTGCAGA	AGCACAAATG
24451	GCCCCAGCCC	TGGGCTGTGT	ACTATCAGCT	CTTTCAAGTC	CTTCCTGGGG
24501	ACCGGGGTGA	GCAGGGATGG	GTTGGAGCTC	AGGATAGGGG	GCAGCGGGGA
24551	GGCGAGCAGA	CTGACCACGC	CCAAGGATGG	AGCCCAAGGT	TACCCGGGTT
24601	CACAGGGCTG	TGAGGTGCTT	CAGGCAGAGA	GTAGGGGTAA	GATAATCAGT
24651	GGAGGTAAGA	GGACATAAAA	TACCTGTAAC	CCAACGATGT	AGGGTCATGA
24701	GATTGTCTTG	GCTCAGTGTG	AGAGAGAGGT	ACCAAAGGTC	ATCTTCCTAA
24751	AATTTAAAG	ACAATAAGAT	TGTCCAGGGT	CCGGCCAGGC	GCAGTGGCTC
24801	ATGTCTGTAA	TCCCAGCACT	TTGGGAGGTC	AAGCTGGGCG	GATCACTTGA
24851	GGTCAGGAGT	TTGAGACCAG	GCTGACCAAC	GTGGTGAAAC	CCCGTCTCTA
24901	CGAAACATAC	AAAAATTAGT	CGGGTGTGGT	GGCACACTCC	TGTAGTCCCA
24951	GCTACTCAGG	AGGCTGAGGC	AGGAGAATAA	TTGCTTGAAT	CTGGGAGGCG
25001	GAGTTGCAG	TGAGCCGAGA	TCATACCACT	GCACTTCAGC	CTGGGCAGCA
25051	GAGCGAGACT	CTGTTTTAAA	AAAAAAAAAA	AAAAAAAGAC	TGTCCACGGA
25101	CAAGTGACAG	AAGGGAGTGT	TTCTGACCTT	CAATTTGTAG	GATGGGCTGG
25151	GCATGGTGGC	TCACAACTGT	AATCCCAGCA	CTTTGGGAGG	TCAAGGTGGG
25201	TGGATTGTCT	GAGCTCAGGA	GTTTGAGACC	AGCCTGGGCA	ACATGAGGAG
25251	ACCCCATCTA	TACAAAAAAT	AGAGAAAATTG	GCTGGGTGCG	GTGGCTCAAC
25301	GCCTGTAATC	CCGGCACTTT	GGGAGGCCAA	AGCGGGTGGA	TCACTTGAGG
25351	TCAGGAGTTC	GAGACCAGCC	TGGCCAACAT	GGTGAAGCCC	CGTCTCTACT
25401	AAAAATACAA	AAAAATTAGC	TGGGCATGGT	GGCACATGCC	TGTAGTCTCA
25451	GCTACTCGGG	AGGCTGAGGC	AGAAGAATCG	CTTGAACCCA	GGAGGCGGAG
25501	GTTGCAGTGA	GCCGAGATCG	CACCACTGCA	CTCCAGCCTG	GCGACTGAGC
25551	AAGACTCTGT	CTCAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAGCCAAT
25601	ATATATATAT	ATATATAGAG	AGAGAGAGAG	AGAGAGAGAG	AGAGAGAGAG
25651	AGATTAGCTG	AGCATGGTGG	CATGTGCCTG	TATTCCCAAC	TCCAACACT
25701	GGGGAGGCTG	AGGTGGGAGG	ATCACTTGAG	CCTAGGAGGT	GGAGGCTGCA
25751	GCGAGCTGAG	ATCACGCCAC	TGCACTCCAG	CCTGGGTGAC	AGAGCAAGAC
25801	CCTGTCTCAA	TTAAAAAATA	AAAAGGGGGC	CGGGCATGGT	GGCTCACGCC
25851	TGTAATCCCA	GCACTTTGGG	AGGCCGAGGC	GGGTGGATCA	CGAAGTCAAG
25901	AGATCGAGAC	CATCCTGGCC	AACATGGTGA	AGCCTCGTCT	CTACTAAAAA
25951	TACAAAAAAT	TAGCCAGGCA	TGGTGGCAGG	CGCCTATAGT	CCCAGCTACT
26001	CAGGAGGCTG	AGGCAGAAGA	ATCACTTGAA	CCCAGGAGGT	GAAGGTTGCA
26051	GTGAGCCAAG	ATTGCGCCAC	TGCACTCCAG	CCTGGCGACA	GAGTGAGACT
26101	CCGTCTCAAA	AAAAAAGAAA	AAATAGATTG	TCTAGGGTCG	AGTGAGAGAA
26151	GGGAGTGTAG	AAGTTTGTCT	GATCTTAAGT	TTGTAGCATC	ATGAGATTGT
26201	TCAGGCTCAA	CCTGATGGGA	TGGGAGACTA	AAGGGCATCT	GGGCTTAGAT
26251	TTGTGAGAAC	TAAGTTTGT	CACCACTGGG	ACCCTGAAAT	TATCTGAACT
26301	TGGGACGGGA	GAGAGGCAAA	TGGATAGCCG	CGGAAGCATG	AGATTGTCCT
26351	GTCTGACAGG	GAGAAGCAAG	GGATTGAGCG	TATTCACGCT	GAAGTACATG
26401	GCATGAGGTT	GGCTGGATAT	TAGGAAAAGGA	TGCTTGTGGT	TGTTCAAGTG
26451	TTGAGTGTGA	GGCCACAAGC	TCGTGCAGGC	TGGAAGTGGG	AAGTTATTCA
26501	AGTTCATGGT	GACAGCAGCA	TGGGATTGGC	TGGGAGTGGT	TGTGGGGGAG

FIGURE 31

26551	GGGTAGGGTG	AGCAGGAAGT	TGTTTGGCGG	GGGTGGTCT	AGGTGGTCT
26601	AAGTTTGCCC	AAACTTTTAC	TGCAGTTTGT	CGTTTGTGTT	TGTTTTTGT
26651	TTTTTTTTTT	TGAGATGGAG	TCTCGTGCTG	TTGCCCAGGC	TGGAGTGCAA
26701	TGGCAGGATC	TCGGCTCACC	GCAACATCCA	CCTCCTGGGT	TCAAGAGATT
26751	CTCCTGCCTC	AGCCTTCCGA	GTAGCTGGGA	TTACAGGCAT	GTGCCACCAT
26801	GCCTAGCTCA	TTTTTTTGGT	ATTTTATAGTA	AAGACGGGGT	TTCACCATGT
26851	TGGCCAGACT	GGTCTCGAAT	TCCTGACCTC	AAGTGATCCA	CCCACTTCGG
26901	CCTCCCAAAG	CGCTGGGATT	ACAGGCACGA	GGCATCGCGC	CCGCCAGTT
26951	TGCTCAAAC	TTTACTGCAG	GTTGCCTTGT	CTCTATGGTG	AGGGGGAGAA
27001	TATTAGGAGG	TTGCCCAGGC	TTATGATAAG	GGAAGGCATG	AGGTGGTGCA
27051	AGTTTTCAAG	TGAGAAATCG	TCCAGGTTCC	CAGTGACAGC	AGAATGAGCT
27101	TGGCTTGCCA	GTAGCTGCAG	AGGGACCCAT	GGCTGTTTCA	GTTTCGCGGT
27151	GAGTGGCAGG	AGGCTCCCGG	GTCCTCTGTG	GGGTGACAC	AAGGTTGTGA
27201	GGGCCTATTA	CCACCATCTC	CACTCCTGAC	CAGTGCTCCT	GATTCTGAC
27251	AATGCCAACG	TCTTCTATGC	CATGAGTCCA	GTCGCCCCCA	GAGACTTCAT
27301	GCTGCGGCGG	AAAGAGGGGA	CCCGGAACAC	TCTGTCTGTC	TCCCCAAGCT
27351	GAGGCAGCCC	TGTCCTCTCC	ACAAGACACA	AGTCCCACAG	GCAAGCTTGC
27401	GACTCTTCTC	CTGGAAAGCT	GCCATCCCCC	AGTAGAGGCC	ACTGTGCTGT
27451	GTATCCCAAG	ACCACCAACC	AACGTAGGCC	CATTGGACCC	CATCTCTTTT
27501	TCTGACTCTG	TTGGTACTAG	ATCCATATTC	CAAAGACATC	AGCCCATGGG
27551	TGGCTGGTGG	AGAGCTCAAT	CCCATAAATG	TAGAAAGAGG	TGGGGCATGG
27601	ATACGTCAAA	TCCCTCCCCA	GAGAAATCTT	ATAAATGTTA	GAGACGCATC
27651	AGAAGTGACA	GATGCGGATG	AAAATAGTGA	CCAGAGTTAT	GAAACAGGTG
27701	TCAGTCTTGT	TTATTTTTCG	CCTGTGTGCC	ATGTTACCC	TTTATCAAGA
27751	TAAAGGAAAA	CAGCTACCAC	ACACACACCC	ACACACACAC	ACACAAACAC
27801	ACAGAGAGAG	AGAAACCTAA	GAGCCAAGAC	CAGCCCGGGC	AACATAACGA
27851	GATCCTGTCT	CTACAAAAAA	TACAAAAATT	TGGCTGGGCG	TGGTGACTCA
27901	CGCCTGTAAT	CCCAGCGTTT	TGGGAGGCCA	AGGCAGGCAG	ATCGCCTGAG
27951	GTCAAGAGTT	CGAGACCAGC	CTGGCCGACA	TGGCAAAACC	CCATCTTCTA
28001	AAAGTACAAA	AAATTAGCCG	GGCGTGGTGT	CATGCACCTG	TAATCCCAGC
28051	TACTGGGGAG	GCTGAGGCAG	GAGAATTGCT	TGAACCCGGA	AGGTGGAGGT
28101	TGCAGTGAGT	GGAAATCACA	CCACTGTACT	CCAGCCTGGG	TGACAGAGCA
28151	AGACCCTATC	TCAAAAACAA	ACAAACAAAC	AAATGAACAA	ACAAAAAATT
28201	TTCTGAGTGT	GGTGATATGA	GACTGTAATC	CTACCTACTT	GGGAGGCTGA
28251	GCTGGGAGAA	TCACCAGAGC	CCTGGGAGGT	TGAAGCTGCA	GTGAGCAGTG
28301	ACTGGGCCCC	TGCACTCCAA	CCTGGAGGAC	AGAGGGAGAT	CCTGTCTCAA
28351	AAACAAAAAA	ACTAAGAGCC	CTAAGAAAGG	TGTTGAGTCG	GGTATGACAC
28401	TCAACCCAGA	TGCCAGAGAG	GATCCTGTCT	GGCCGGACAC	AGTGGCTCAG
28451	GAGGGTAATC	CCAGTACTTT	GGGAGGCTGA	GGTAAGAGGA	TTGCTTAAGT
28501	TCAGGATGCC	GAGAGCAGCC	TGGGCAATAC	AGTGAGATCA	CATCTCACTA
28551	AATAAATAAA	TAAAGGATCC	TATCACACAA	AGAGGGTTTA	GGACTTCCTT
28601	CCCCAACATT	TTTGGGGTGA	TATGCCTCTT	TTCTACTGTA	TATATGGGAG
28651	AGTGACTAAC	TGAAATTCCA	TCAGAATTAG	AAACAAATAG	CATCATTACC
28701	CATGAGTCAA	TAAGGGCTGT	GAGGATGGGC	CCTTTCACCT	GCCCTCACCT
28751	TCTTCCTCTT	CCTGTCACAG	ATAACCCATC	TGTGCAAAGA	AGAGAAAAAG
28801	AGGTTGGGTG	TGGTGGCTCA	CATCTGTAAT	CCCAGCACTT	TGGGAGGCTA
28851	AGGTGGAAGG	ATTTTGAGCC	CAGGAGTTTG	AGACCAGCCT	GGGCAACATA
28901	GTGAGACCCC	ATTTCTACAA	AAAAATACAA	AGATTGGCCA	GGCGCGGTGG
28951	CTCACGCCTG	TAATCCCAGC	ACTTGGGAGG	CTGAGGCAGG	CGGATCATGA
29001	GGTCAGGAGA	TCGAGACCAT	CCTGGCTAGG	TGAAACCCCG	TCTCTACTAA
29051	AAATGCAAAA	AAATTAGCCG	GGCGTGGTGG	CGGGCGCCTG	TAGTCCCAGC
29101	TACTCGGGAG	GCTGAGGCAG	GAGAATGGCG	TGAACCCAGG	AGGCGGAGCT
29151	TGCAGCGAGC	CGAGATCGCA	CCACTGCACT	CCAGTCTGGG	CGACAGAGGG
29201	AGACTCCATC	TCAAAAAATA	AAATAAAATT	TAGCCAGGTT	TGGTGTCTCTG
29251	CACCTGTAGT	CTCAGCTACT	CTGGAGGCTG	AAGCACGAGG	ATCACTTGAG
29301	CCCAAGAGGT	GGAGGTTGCA	GTGAGCCGAG	ATTACTGCAC	TCCAGCCTGG
29351	GTGACAGAGC	GAGATCCTGT	TTCAAAAAGC	AAAAAAAAGG	GCCAGGCGCA
29401	GTGCTCACAC	CTGTAATCCC	ATCATTTTGG	GAGGCTGAGG	TGGGCGGATC
29451	ACTTGAGGTC	AGGAGTTCAA	GGTCAGCCTG	GCCAACATGG	TAAAACCTCG

FIGURE 3J

```

29501 TCCCTACTAA AAAATATAAA AATTAGCTGG GCATGGTGGT GGGTGCCTGT
29551 AATCCCAGTT ACTCAGGAGG CTGAGGCAGG AGAATTGCTT GAATCCAGAA
29601 GGTGGAGGTT GCAGTGAACC GAGATCATGC CATTGCACTC CAGCCTGCGT
29651 GACAAAGTGA GACTGTATCT CAAAAAAAAA AAAAAAATG CTGGGCACAG
29701 TGGCTCTAGC ACTTTGGGGG GGCAAGACGG GTGGATTGCT TGAGGCCAGG
29751 ATTCCAAAAC CAGCCTGGCC AACATGGTGA AACCCTTCT CTTAAATA
29801 TACAAAAAAT TAGCCGGGCA TGGTGGCAGG CTCTTGTAA CTCCAGCTACT
29851 CGGTAGGCTG AGGCAGGATA ATCACCTGAA CCAGGCAGGC AGAGGTTGCA
29901 GTGAGTCGAG ATCGCTCCAC TGCACTCCGG CCTGGGCAAC AAGAGCAAAA
29951 TTCTGTCTGG AAAAAAAAAA AAAAGAAAAA GAAAAGGATT GTGAGGATGA
30001 AAAGAGAGGC GTGAGCTCTC TGTCAGCGTT GGAGTACAAT AGAGAGGATG
30051 AAATGAGCTG TAGGGCGAAC TGCTACATAG TCACAACCAC AATAATATGC
30101 CCACTTATGA GCTCCTACTC AGCAGAGAAC ATCAGCTATG GTCTTTACAT
30151 CTCATTGCAC TAATCGAGTT CTTTCTGTTG CAAGCGACCA AAAACCCAAT
30201 TCAAAGAGGC ATGTGCAAAA AAGGACATTT GTGGCTTATG CAGTTGAAAT
30251 GTCCAATGAG TAGGGCTTCA GGCACAGTTG CATCCAGGCA CTCATAAGAT
30301 GTCATCAGGG TTTCTTGCT GTCTCTTTGC TCTGATTGCT TCTGAGAATG

```

(SEQ ID NO:3)

FEATURES:

```

Start: 3000
Exon: 3000-3032
Intron: 3033-3493
Exon: 3494-3607
Intron: 3608-5219
Exon: 5220-5443
Intron: 5444-5611
Exon: 5612-5665
Intron: 5666-6130
Exon: 6131-6342
Intron: 6343-15419
Exon: 15420-15562
Intron: 15563-15649
Exon: 15650-15865
Intron: 15866-16855
Exon: 16856-16959
Intron: 16960-17050
Exon: 17051-17135
Intron: 17136-19043
Exon: 19044-19100
Intron: 19101-19208
Exon: 19209-19295
Intron: 19296-19502
Exon: 19503-19535
Intron: 19536-19617
Exon: 19618-19739
Intron: 19740-21452
Exon: 21453-21548
Intron: 21549-21641
Exon: 21642-21710
Intron: 21711-21798
Exon: 21799-21895
Intron: 21896-24152
Exon: 24153-24305
Intron: 24306-24391
Exon: 24392-24506
Intron: 24507-27233
Exon: 27234-27349
Stop: 27350

```

FIGURE 3K

ALLELIC VARIANTS (SNPs):

DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
1695	A	C G	Beyond ORF(5')			
4607	T	C	Intron			
5133	G	C	Intron			
6190	A	C	Exon	162	H	P
6196	T	C	Exon	164	V	A
7403	T	G C	Intron			
9981	C	T	Intron			
9998	-	T	Intron			
11050	T	C	Intron			
11772	T	A G	Intron			
12797	T	-	Intron			
12801	T	-	Intron			
12857	T	C	Intron			
17525	C	T	Intron			
17552	G	A	Intron			
17572	T	G	Intron			
18846	G	A	Intron			
18871	T	C	Intron			
18872	C	A	Intron			
19600	T	C	Intron			
21013	A	G	Intron			
22055	A	G	Intron			
22112	T	C	Intron			
22113	T	C	Intron			
22221	C	T	Intron			
22360	T	C	Intron			
22543	G	A	Intron			
23751	A	T	Intron			
23764	C	T	Intron			
23782	G	A	Intron			
24593	C	T	Intron			
24673	C	T	Intron			
24730	T	C	Intron			
24750	A	G	Intron			
24764	A	G	Intron			
24780	A	T	Intron			
24856	A	G	Intron			
26136	A	G	Intron			
26331	C	G	Intron			
28362	A	G C	Beyond ORF(3')			
28745	T	C	Beyond ORF(3')			

Context:

DNA

Position

1695 ATCACTCAAGATCCGTTCCAGAAATTAATCGAAAGTCACAAGAAGCACCGTCGCTCTCG
GAGGTCCTAGACTCGTCCTGACACCCACCAGGCGGCCCTTCGGAGCCCCGAATCTCCG
GGTCTAGCGCACGCCACGGGCGCTTCAGGGGCTGAACGCGGCCGACCGGGAACGGAGGC
GGCCAGCGGCGCCCGGAGGGGAGGAAGGGGCGGGCCAGACGTTCCACAGTAAATCTCC
CCAGCTGGGTCCGCCCCGGCCTCAGAGTTGCGCAATAAATGTTACCGACCATGCCCTGG
[A,C,G]
TATTCATCTGTTTTTGACCCTGCACCACCCAAGAGACGGCTGTCCCTGAAAACCCAGGGC
CACAGACTGCCTCCTCCAACCTGGGTCATGATGACTCCCATCAGCTAGTGACGCAGATGG

FIGURE 3L

AGCTTAAAAATGGGAGATGGCCCGATGTAGTGGTTTTATGCCTGTAATCCCAGTATTTTG
GGAGGCTGAGTTGGGAGGATCACTTGAGTCCAGGAGCTCCAGGCTGCAGTGAGCTATGAT
CGTGCTACTGCACTCCAGCCTGGGCCACAGAGCCAGACCCTGTCTCAATAAATAAAATAA

4607 CCCAGCAGTGGTCCCAGCTACTCAGGAGGCTGGGGCAGAGGGACCGCTTGAGTCTAGGAC
TTGGAGGCTGCAGTGAGCTATGATTGTGCCACTGCACCCAGCCTGGGCAACAAAAACAAG
ACCTGTTTTCTAAAAAAACAAACCAAAACAATAACTCCAAGAAGCCGGGAGACAGAGGAA
TCACATGAAAGAATGGTGCTACAGGCGGGGCGAGGTGGCTCACGCCTGTAATCCCAGCAC
TTTGGGAGGCCGAGGCAAGTGGATCATCAGGTCAGGAGTTCAAGACCAACCTGGCCAAGA
[T, C]
GGTGAAACACCGTCTCTACTAAAAATACGAAAAAACTAGCTGGGCTTGGTGGCGGGTGCC
TGTAAATCCCAGATACTTGGGAGGCTGAGGCAGAGAATTGCTTGAACCCAGGAGGCGGAGG
TTGCAGTGAGCCAAGATCAGGCCACTGAACTCCAGCCTAGGTGACAGAGTGAGATTCTGT
CTCAAAAAAAAAAAAAAAAAAAGTGGTGCTAGGGGCTGGGCACGGTGGTTACGCCTGTAA
TCCTAGCCCTTTGGGAGACTTTGGGAGGCCAAGGGGGGCAGATTACTTGAGGTCAGGAGT

5133 GGTTCACGCCTGTAATCCTAGCCCTTTGGGAGACTTTGGGAGGCCAAGGGGGGCAGATTA
CTTGAGGTGAGGAGTTCGAGACCAGTCTGACCAACATGGTGAAACCTATCTCTACAAAA
ATACAAAAATTAGCTGGGCTTGGTGGTGTGCGCCTGTAGTTTCAGCTACTTGGAGGCTGA
GGAAGGAGGATTGCTTGAACCCAGGAGGCAGAAGTTGAAGTGACCCAAGATCGTGCCACT
GCACTCCAGCCTGGGCAACAGAGTGAGACTCTGTCTCAAAAAAAAAAACAAAAAAAAAA
[G, C]
AGTGGTGCTAGTGATGAATGTGACTAGAGAAGGGGTGCTGTGAGGACCACTCCTGCTCTC
TCATGGCCACCTCTCCCTCCTGCAAGGCTCCCAGCCCCATTGCCAATACCTTCCTCCACT
ATCGAACCAGCAAGGTGAGGGTGCTGAGGGCAGCGCGCTGGAGCGGCTGGTGGGAGAGT
TGGTGTGTTGGAGACCGTGAGCAGGACCCAGCTTCATGCCCGCCTTCCTGGCCACCTACC
GGACCTTTGTACCACTGCCTGCCTGCTGGGCTTTCTGCTGCCACCAATGCCACCGCCCC

6190 CTGTCTCTACTAAAAATACAACAACAACAAAAAAGGTAGCTGGGTGTGGTGGCGCATGC
CTGTAGTCCCAGCTACTCGGGAGGCTGAGGTTGCAGTGAGTCAAGATCAGGCCATTGCAC
TGCAGCCTTGGTGACACAGTAAGACTCTATCTCAAAAAAAAAAAAAAAAAAAGGTACCA
GGAGTCATATTCTATGTCCCCACTCTGGACCCAGCTCTGAGACCCTGCCTCTCTGGCCA
GGGCTGTGGTGTGAGTGCTGGGCTCCTGGCTGCAGGACCACCCTCAGGATTTCCGAGACC
[A, C]
CCCTGTCCATTGCGACCTGGGCAGTGTCCGAACCTTTCTGGGCTGGGCGGCCCCAGGGAG
TGCTGAGGCTCAAAAAGCAGAGAAGCTTCTGGAAGATTTTTTGGAGGAGGCTGAGCGAGA
GCAGGAAGAGGAGCCGCTCAGGTGTGGACAGGTGAGGGGTTTTTTCAGATCCAGTCGTGTT
CTGAGAAGGCCTTTCTGTCTGCTTCTTCCCACACAGGCTTTCTCTCCCTCTCAGAGCT
ACAAAACCTTAAGCAAGATTTTAACTCTAAGCCTCAATTTCTTCATCTTTACAATGGGGA

6196 CTACTAAAAATACAACAACAACAAAAAAGGTAGCTGGGTGTGGTGGCGCATGCCTGTAG
TCCCAGCTACTCGGGAGGCTGAGGTTGCAGTGAGTCAAGATCAGGCCATTGCACTGCAGC
CTTGGTGACACAGTAAGACTCTATCTCAAAAAAAAAAAAAAAAAAAGGTACCAGGAGTC
ATATTCTATGTCCCCACTCTGGACCCAGCTCTGAGACCCTGCCTCTCTGGCCAGGGCTG
TGGTGTGAGTGCTGGGCTCCTGGCTGCAGGACCACCCTCAGGATTTCCGAGACCACCCTG
[T, C]
CCATTGCGACCTGGGCAGTGTCCGAACCTTTCTGGGCTGGGCGGCCCCAGGGAGTGCTGA
GGCTCAAAAAGCAGAGAAGCTTCTGGAAGATTTTTTGGAGGAGGCTGAGCGAGAGCAGGA
AGAGGAGCCGCTCAGGTGTGGACAGGTGAGGGGTTTTTTCAGATCCAGTCGTGTTCTGAGA
AGGCCTTTCTGTCTGCTTCTTCCCACACAGGCTTTCTCTCCCTCTCAGAGCTACAAAA
CTTAAGCAAGATTTTAACTCTAAGCCTCAATTTCTTCATCTTTACAATGGGGATAATAA

7403 TTAATTTAATTTAATTTAATTTTATTTTTTTATCTATTTTTTTTTTAGACAGAGTCTC
GCACTGTTACCCAGGCTGGAGTGCACTGGCACAATCTTGGCTCACTGCAACCTCCACCTC
CTGGGCTCAAGCCATTCTCCTGTCTCAGCCTCCTGAGTAGCTGGGACCACAGGCGCATGT
CACCACGCCCCGCTAATTTTTTTGTAAAGGTGAGGTGTGCCATGTTGCCAGGCTGGTC
TCAAACCTCTGAACCTCAAGTGATCTGCCTGCCTTGGCCTCCCAAAATGCTGGGATTACAG
[T, G, C]
CATAAGCCATTGTGCATGCGTAGCCTCCTTACTTGATTATTGGCTTTTGCTCATCTCATA

FIGURE 3M

9981

9998

11050

11772

12797

FIGURE 3N

GGACTACAGGCGCCCGCCACCACGCCTGGCTAATTTTTCTATTTTTAGTAGAGACGGG
GTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCTACCCGCCTTGG
CCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCCCCGGCCTTCTTTCTTTCT

12801 GTAGAGTGTACGGTGCTGTTGTA CTGACAGCAACAAGCACCAACGAACGCACAGGAGGG
CACTGGTGAGGCAAAGACAGCAACATAGGTTCTGGGGACATCATTTTCCAACTTGTGAA
CAACATTTGCAATTTGCAAAACAAAACAAAGCCAGACTTTCGTGGTCTTGCATTCTTGG
AGCCAAAAAATTTGTGTTTATGAACAAAATAGTCAGGTTCTAGGTGCATATTATTGCAA
ACATGTTTTCTTTCTTTTGTTTTGTTTGTTTGTTTGTTTGTTTGTTTGTTTGT
[T,-]
GAGATGGAGTCTCGCTCTGTGCGCCAGGCTGGAGTGAGTGGCATGATCTCGGCTTACTG
CAAGCTCCGCTCGCCGTTACGCCATTCTCCTGCCTCAGCCTCCTGGGTAGCTGGGAC
TACAGGCGCCCGCCACCACGCCTGGCTAATTTTTCTATTTTTAGTAGAGACGGGTTT
CACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCTACCCGCCTTGGCCTC
CCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCCCCGGCCTTCTTTCTTTCTTTT

12857 AGGGCACTGGTGAGGCAAAGACAGCAACATAGGTTCTGGGGACATCATTTTCCAACTTG
TGAACAACATTTGCAATTTGCAAAACAAAACAAAGCCAGACTTTCGTGGTCTTGCATTCT
TTGGAGCCAAAAAATTTGTGTTTATGAACAAAATAGTCAGGTTCTAGGTGCATATTATT
GCAAACATGTTTTCTTTCTTTTGTTTTGTTTGTTTGTTTGTTTGTTTGTTTGT
TTTTTGAGATGGAGTCTCGCTCTGTGCGCCAGGCTGGAGTGAGTGGCATGATCTCGGCT
[T,C]
ACTGCAAGCTCCGCTCGCCGTTACGCCATTCTCCTGCCTCAGCCTCCTGGGTAGCTG
GGACTACAGGCGCCCGCCACCACGCCTGGCTAATTTTTCTATTTTTAGTAGAGACGGG
GTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCTACCCGCCTTGG
CCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCCCCGGCCTTCTTTCTTTCT
TTTTTTTTTTTTTGAGACAAAGTCTCTGTACCCAGGCTAGAGTGCCGTGGCGTGAC

17525 TCAGGACACTCTGTCTTCCCTACCGCTCAGCAATGACCTTATCCTTGTCCCTGGCGGGT
TGACGTTTTCTTCTCTACTTCTGCGTTATAGTTGACTGTGAGTGACTGCCCTATT
TATTCACCTCAGCAAAACACAAGAAGTCACAAAGAAAAGTTACTTAAGGCCAGAGTCATA
GCACAGGTTGGGAACAAAAAATGTTCTGAGGACTTTACCTTGATAAGCAAACTAAAA
AATGTGTGTCAAAGTCTGGCTTATTTATAGGCAAGATTTAGATTCTCATTGCAATCAGG
[C,T]
GCTGGTTTTAGAGTGAATCTAGAATGGATCCCTGGGCCTGGAACATTCTCCACCCCTCC
AGGTTTTGCATGCAACTTGCTCACTCACCTCCTTCTGGTCTCTGATTAAATGTCCCTGCCT
CTGAGAGGCCTTCCCAGCCTCCATCATCCCCAAAACCACACATCTGGTTTTTGTGTGTG
TTGTTGT
CTCGCTCTGTACCCAGACTGGAGTGAGTGGCACGATCTTGGCTCACTGCAACCTCCAC

17552 TCAGCAATGACCTTATCCTTGTCCCTGGCGGGTGCACGTTTTCTTCTCTACTTCTCT
GCGTTATAGTTGACTGTGAGTGACTGCCCTATTTATTCACCTCAGCAAAACACAAGAAGTC
ACAAAGAAAAGTTACTTAAGGCCAGAGTCATAGCACAGGGTGGGAACAAAAAATGTT
CTGAGGACTTTACCTTGATAAGCAAACTAAAAAATGTGTGTCAAAGTCTGGCTTATTT
ATAGGCAAGATTTAGATTCTCATTGCAATCAGGCGCTGGTTTTTAGAGTGAATCTAGAAT
[G,A]
GATCCCTGGGCCTGGAACATTCTCCACCCCTCCAGGTTTGCATGCAACTTGCTCACTCAC
CTCCTTCTGGTCTCTGATTAAATGTCCCTGCCTCTGAGAGGCCTTCCCAGCCTCCATCAT
CCCCAAAACCACACATCTGGTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
TTGTTCTTTGT
AGTGGCACGATCTTGGCTCACTGCAACCTCCACCTCCCAGGATCAAGCAATTCTCTCTGC

17572 GTCCCTGGCGGGTGCACGTTTTCTTCTCTACTTCTGCGTTATAGTTGACTGTGAG
TGACTGCCCTATTTATTCACCTCAGCAAAACACAAGAAGTCACAAAGAAAAGTTACTTAA
GGCCAGAGTCATAGCACAGGGTGGGAACAAAAAATGTTCTGAGGACTTTACCTTGATA
AGCAAACTAAAAAATGTGTGTCAAAGTCTGGCTTATTTATAGGCAAGATTTAGATTCT
CATTGCAATCAGGCGCTGGTTTTTAGAGTGAATCTAGAATGGATCCCTGGGCCTGGAACA
[T,G]
TCTCCACCCCTCCAGGTTTGCATGCAACTTGCTCACTCACCTCCTTCTGGTCTCTGATTA

FIGURE 30

ATTACAGGCGGAGCTACACACCTGGCTAATTTTTTTGTATTTTAGTAGAGACGGGG
TTTTACTATATTGGCCAGGCTGGTTTCGAACCTCTGACCCGAAGCAATCCTCTCGCCTCA
GCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACTGCACCCAGTCTGACCCCTGACTT
TAATCCTGACCCAATTTGATTCTTAGTGCCACCCTGTGAATCTCTTTGTGACCTCCTGA
CCAGCCATCCTGTCCCATCTCTGATAAGACCTTGATGCTCAATGACCCTCATTTACCACC

22055 GAGTCAGCACAGCCACCCACCTCAGCCTCTGCATCTCCCCAGTGTGTCCCCAGGGTCA
CCCCCTCAAGTCCTAGAAGCAGAGATGCTCCTGCTGGCAGTCCCCCGGCCTCTCCAGGG
CCCCAGGGCCCCAGCACCAAGGTACCAAGACGGCTTGTGTGTGCATGCGGGCCTGCGGGC
ACCCAGGCTCTGTGTGTGTGCACGTGTGTGTGCATGCACATGTGTACACACAGGATTGTG
GGGCCAGGAGTGTATACAGGAGGCACACTGAGCGCCCGGGGTATCCATCCAGGGGATTGC
[A, G]
TGCATCTGCACGGCCCTGTTTGGGTGATCACTCATAAATCCGACTCGTGCTCAGATTTGG
ACCTGTGTAAGTCTTGCCCATGGGTCACTAGGGTGCAATCACATCACACCCCTTTTFA
TTTGAAACAGGGTCTTCTTGCTCTGTACCCAGGCTGAAGTGACGCGGTGCAATCTCAGC
TCACCGCAACTTCCACCCCTCCCCAGGCTCAAGCAATCCTTCCACCTCAGCCTCCCAAG
TAGCTAGGACCACAGGTGTGCACCACCATGCCCTGCTATTTTTTTATTTAGTAGAGATG

22112 TCACCCCTCAAGTCCTAGAAGCAGAGATGCTCCTGCTGGCAGTCCCCCGGCCTCTCCA
GGGCCCCAGGGCCCCAGCACCAAGGTACCAAGACGGCTTGTGTGTGCATGCGGGCCTGCG
GGCACCCAGGCTCTGTGTGTGTGCACGTGTGTGTGCATGCACATGTGTACACACAGGATT
GTGGGGCCAGGAGTGTATACAGGAGGCACACTGAGCGCCCGGGGTATCCATCCAGGGGAT
TGCATGCATCTGCACGGCCCTGTTTGGGTGATCACTCATAAATCCGACTCGTGCTCAGAT
[T, C]
TGGACCTGTGTAAGTCTTGCCCATGGGTCACTAGGGTGCAATCACATCACACCCCTTT
TTATTTGAAACAGGGTCTTCTTGCTCTGTACCCAGGCTGAAGTGACGCGGTGCAATCTC
AGCTCACCGCAACTTCCACCCCTCCCCAGGCTCAAGCAATCCTTCCACCTCAGCCTCCC
AAGTAGCTAGGACCACAGGTGTGCACCACCATGCCCTGCTATTTTTTTATTTAGTAGAG
ATGAGGTTTCGCCATGTTGCCCAGGTGGGTTTCGAACCTCTGAGCTCAAACAATGCACTC

22113 TCCTGCTGGCAGTCCCCCGGCCTCTCCAGGGCCCCAGGGCCCCAGCACCAAGGTACCAAG
ACGGCTTGTGTGTGCATGCGGGCCTGCGGGCACCCAGGCTCTGTGTGTGTGCACGTGTGT
GTGCATGCACATGTGTACACACAGGATTGTGGGGCCAGGAGTGTATACAGGAGGCACACT
GAGCGCCCGGGGTATCCATCCAGGGGATTGCATGCATCTGCACGGCCCTGTTTGGGTGAT
CACTCATAAATCCGACTCGTGCTCAGATT
[T, C]
GGACCTGTGTAAGTCTTGCCCATGGGTCACTAGGGTGCAATCACATCACACCCCTTTT
TATTTGAAACAGGGTCTTCTTGCTCTGTACCCAGGCTGAAGTGACGCGGTGCAATCTCA
GCTCACCGCAACTTCCACCCCTCCCCAGGCTCAAGCAATCCTTCCACCTCAGCCTCCCA
AGTAGCTAGGACCACAGGTGTGCACCACCATGCCCTGCTATTTTTTTATTTAGTAGAGA
TGAGGTTTCGCCATGTTGCCCAGGTGGGT

22221 GCGGGCTGCGGGCACCCAGGCTCTGTGTGTGTGCACGTGTGTGTGCATGCACATGTGTA
CACACAGGATTGTGGGGCCAGGAGTGTATACAGGAGGCACACTGAGCGCCCGGGGTATCC
ATCCAGGGGATTGCATGCATCTGCACGGCCCTGTTTGGGTGATCACTCATAAATCCGACT
CGTGCTCAGATTTGGACCTGTGTAAGTCTTGCCCATGGGTCACTAGGGTGCAATCACA
TCACACCCCTTTTTATTTGAAACAGGGTCTTCTTGCTCTGTACCCAGGCTGAAGTGACAG
[C, T]
GGTGCAATCTCAGCTCACCGCAACTTCCACCCCTCCCCAGGCTCAAGCAATCCTTCCAC
CTCAGCCTCCCAAGTAGCTAGGACCACAGGTGTGCACCACCATGCCCTGCTATTTTTTTT
ATTTAGTAGAGATGAGGTTTCGCCATGTTGCCCAGGTGGGTTTCGAACCTCTGAGCTCAA
ACAACTGCCTCACCTCGGCCTCCCAAGTGCTGGGATTACAGGTGTGAGCCACCCGACCC
AGCCTACACTTTTTTGAGGACATGTATGTCCCTAAGAATCTGCATACCATGGCAGACACG

22360 TCTGCACGGCCCTGTTTGGGTGATCACTCATAAATCCGACTCGTGCTCAGATTTGGACCT
GTGTAAGTCTTGCCCATGGGTCACTAGGGTGCAATCACATCACACCCCTTTTTATTTG
AAACAGGGTCTTCTTGCTCTGTACCCAGGCTGAAGTGACGCGGTGCAATCTCAGCTCAC
CGCAACTTCCACCCCTCCCCAGGCTCAAGCAATCCTTCCACCTCAGCCTCCCAAGTAGC
TAGGACCACAGGTGTGCACCACCATGCCCTGCTATTTTTTTATTTAGTAGAGATGAGGT

FIGURE 3Q

[T,C]
TCGCCATGTTGCCCAGGTGGGTTTCGAACTCCTGAGCTCAAACAATGCACTCACCTCGGC
CTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCGCACCCAGCCTACACTTTTTTTGAGG
ACATGTATGTCCCTAAGAATCTGCATACCATGGCAGACACGGTGGCTATTGCCTGTGATC
CCGGCACTTTGGGAAGCCAAAGTGGGAGGATTGCTTGAGGCCGGGAGTTCAAGACCAGCC
TGGGCAACATAGTGAGACCCTATTTCTATTAAGTCAAAAAAATTAGCTGGGTGTAGTC

22543 AACCTCCACCCCTCCCCAGGCTCAAGCAATCCTCCACCTCAGCCTCCCAAGTAGCTAG
GACCACAGGTGTGCACCACCATGCCCTGCTATTTTTTTTATTTAGTAGAGATGAGGTTTC
GCCATGTTGCCCAGGTGGGTTTCGAACTCCTGAGCTCAAACAATGCACTCACCTCGGCCT
CCCAAAGTGCTGGGATTACAGGTGTGAGCCACCGCACCCAGCCTACACTTTTTTTGAGGAC
ATGTATGTCCCTAAGAATCTGCATACCATGGCAGACACGGTGGCTATTGCCTGTGATCCC
[G,A]
GCACTTTTGGGAAGCCAAAGTGGGAGGATTGCTTGAGGCCGGGAGTTCAAGACCAGCCTGG
GCAACATAGTGAGACCCTATTTCTATTAAGTCAAAAAAATTAGCTGGGTGTAGTCCCA
GCTACTCAGCAGGCTGAGGTGGAAGGATCGCTTGAGTTTGAGGCTGCAGTGAGTACGAT
CATGCCACGGCACTCTAGCCTGCATGATAGAGCGAGATCCTGTTTATGAAGAAAAAGAGA
CTGGGCACGGTGGCTACGCCTGTAATCCAGCACTCTGGGAGGCCGAGGTGGGCGGATC

23751 TTCATGAGATCATGTGTGGGGTTGTTTATTGGCATGGGCTGTGGGTGTATAACCGCTGTC
AGCATATGTATGTACACAGGATTTCTTGATGATGAGCATGGGTGTGTGTATATGGACAC
TGTTTCATGTCTGTTTCTATAACAGGTAACCAAAGTCTGTATATGGTAGGGTGGTGTATAT
GCAGGCTTGTAATGTACTCCAGTTGCATGTCCAGGCTCTGCATGTGTAGGGGGTAGTA
GTATGTTTTCTTGAGATTTTATTTTATTTTATTATTTATTTATTTTATTTTATTTTATTTT
[A,T]
CTTGCTCTGTACGCAGGCTGGAGTGCAGTAGCGTGATCTTGGCTCACTGCAACCTCTGC
CTCTCAAGTTCAAGTGATTCTCCTGCCTCGGCCTCCCAAGTAGCTGGGATTACAGGCATG
CGCCACCAGGCCCTGCTAATTTTTGTATTTTAGTAGAGACGGAGTTTACCACGTTGGC
CAGGCTGGTCTTGAACCTCCGACCTCAAGTGATCCGCCACCTCGGCCTCCCAAAATGCT
GAGATTACAGGCATGAGCCACTGCGCCACGCAATGTTTTCTTGAGATTTTAAATGTGGG

23764 GTGTGGGGTTGTTTATTGGCATGGGCTGTGGGTGTATAACCGCTGTGAGCATATGTATGT
ACACAGGATTTCTTGATGATGAGCATGGGTGTGTGTATATGGACACTGTTTCATGTCTGT
TTCTATAACAGGTAACCAAAGTCTGTATATGGTAGGGTGGTGTATATGCAGGCTTGTA
TGTAATCCAGTTGCATGTCCAGGCTCTGCATGTGTAGGGGGTAGTAGTATGTTTTCTTG
AGATTTTATTTTATTTTATTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTT
[C,T]
GCAGGCTGGAGTGCAGTAGCGTGATCTTGGCTCACTGCAACCTCTGCCTCTCAAGTTCAA
GTGATTTCTCCTGCCTCGGCCTCCCAAGTAGCTGGGATTACAGGCATGCGCCACCAGGCC
TGCTAATTTTTGTATTTTAGTAGAGACGGAGTTTACCACGTTGGCCAGGCTGGTCTTG
AACTCCCGACCTCAAGTGATCCGCCACCTCGGCCTCCCAAAATGCTGAGATTACAGGCA
TGAGCCACTGCGCCACGCAATGTTTTCTTGAGATTTTAAATGTGGGGCTATTGAATGCA

23782 GCATGGGCTGTGGGTGTATAACCGCTGTGAGCATATGTATGTACACAGGATTTCTTGTT
ATGAGCATGGGTGTGTGTATATGGACACTGTTTCATGTCTGTTTCTATAACAGGTAACCA
AAGTCTGTATATGGTAGGGTGGTGTATATGCAGGCTTGTAATGTACTCCAGTTGCATGT
CCCAGGCTCTGCATGTGTAGGGGGTAGTAGTATGTTTTCTTGAGATTTTATTTTATTTTA
TTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
[G,A]
GCGTGATCTTGGCTCACTGCAACCTCTGCCTCTCAAGTTCAAGTGATTCTCCTGCCTCGG
CCTCCCAAGTAGCTGGGATTACAGGCATGCGCCACCAGGCCCTGCTAATTTTTGTATTTT
TAGTAGAGACGGAGTTTACCACGTTGGCCAGGCTGGTCTTGAACCTCCGACCTCAAGTG
ATCCGCCACCTCGGCCTCCCAAAATGCTGAGATTACAGGCATGAGCCACTGCGCCACG
CAATGTTTTCTTGAGATTTTAAATGTGGGGCTATTGAATGCACCAAGTGGTGGCTGGGGTG

24593 TCGAAGCATCTTGGTGAGGGGCTGGGCTGGGGGTCTGCTGGAGGCTGCCCTGCCCTTGGG
GCCGGGGCCCTCACCTCACTCCCCGCCCTCTCTCCAGCTGACCAGTCAGGACAAAGCC
CCCAGCGTGGTCCGGCGAGCCTTGAGAAGCACAATGTGCCCCAGCCCTGGGCTGTGAC
TATCAGCTCTTCAAGTCCTTCTGGGGACCGGGTGAGCAGGGATGGGTGGAGCTCAG

FIGURE 3R

GATAGGGGGCAGCGGGGAGGCGAGCAGACTGACCACGCCCAGGATGGAGCCCAAGGTTA
 [C,T]
 CCGGGTTACAGGGCTGTGAGGTGCTTCAGGCAGAGAGTAGGGTAAGATAATCAGTGGA
 GGTAAAGAGGACATAAAATACCTGTAACCCAACGATGTAGGGTCATGAGATTGTCTTGGCT
 CAGTGTGAGAGAGAGGTACCAAAGGTCATCTTCCTAAAATTTAAAAGACAATAAGATTGT
 CCAGGGTCCGGCCAGGCGCAGTGGCTCATGTCTGTAATCCCAGCACTTTGGGAGGTCAAG
 CTGGGCGGATCACTTGAGGTGAGGAGTTTGAGACCAGGCTGACCAACGTGGTGAAACCCC

24673 TCCCGCCCTCTCTTCAGCTGACCAGTCAGGACAAAGCCCCAGCGTGGTCCGGCGAGC
 CTTGCAGAAGCACAATGTGCCCCAGCCCTGGGCCTGTGACTATCAGCTCTTTCAAGTCCT
 TCCTGGGGACCGGGGTGAGCAGGGATGGGTTGGAGCTCAGGATAGGGGGCAGCGGGGAGG
 CGAGCAGACTGACCACGCCCAGGATGGAGCCCAAGGTTACCCGGGTTACAGGGCTGTG
 AGGTGCTTCAGGCAGAGAGTAGGGTAAGATAATCAGTGGAGGTAAGAGGACATAAAATA
 [C,T]
 CTGTAACCCAACGATGTAGGGTCATGAGATTGTCTTGGCTCAGTGTGAGAGAGAGGTACC
 AAAGGTCATCTTCCTAAAATTTAAAAGACAATAAGATTGTCCAGGGTCCGGCCAGGCGCA
 GTGGCTCATGTCTGTAATCCCAGCACTTTGGGAGGTCAAGCTGGGCGGATCACTTGAGGT
 CAGGAGTTTGAGACCAGGCTGACCAACGTGGTGAAACCCCGTCTCTACGAAACATACAAA
 AATTAGTCGGGTGTGGTGGCACACTCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGG

24730 AGCCTTGCAAGCACAATGTGCCCCAGCCCTGGGCCTGTGACTATCAGCTCTTTCAAGT
 CCTTCCTGGGGACCGGGGTGAGCAGGGATGGGTTGGAGCTCAGGATAGGGGGCAGCGGGG
 AGGCGAGCAGACTGACCACGCCCAGGATGGAGCCCAAGGTTACCCGGGTTACAGGGCT
 GTGAGGTGCTTCAGGCAGAGAGTAGGGGTAAGATAATCAGTGGAGGTAAGAGGACATAAA
 ATACCTGTAACCCAACGATGTAGGGTCATGAGATTGTCTTGGCTCAGTGTGAGAGAGAGG
 [T,C]
 ACCAAAGGTCATCTTCCTAAAATTTAAAAGACAATAAGATTGTCCAGGGTCCGGCCAGGC
 GCAGTGGCTCATGTCTGTAATCCCAGCACTTTGGGAGGTCAAGCTGGGCGGATCACTTGA
 GGTCAAGGATTTGAGACCAGGCTGACCAACGTGGTGAAACCCCGTCTCTACGAAACATAC
 AAAAATTAGTCGGGTGTGGTGGCACACTCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGC
 AGGAGAATAATTGCTTGAATCTGGGAGGCGGAGGTTGCAGTGAGCCGAGATCATACCACT

24750 TGCCCCAGCCCTGGGCCTGTGACTATCAGCTCTTTCAAGTCCTTCCTGGGGACCGGGGTG
 AGCAGGGATGGGTTGGAGCTCAGGATAGGGGGCAGCGGGGAGGCGAGCAGACTGACCACG
 CCCAAGGATGGAGCCCAAGGTTACCCGGGTTACAGGGCTGTGAGGTGCTTCAGGCAGAG
 AGTAGGGTAAGATAATCAGTGGAGGTAAGAGGACATAAAATACCTGTAACCCAACGATG
 TAGGGTCATGAGATTGTCTTGGCTCAGTGTGAGAGAGAGGTACCAAAGGTCATCTTCCTA
 [A,G]
 AATTTAAAAGACAATAAGATTGTCCAGGGTCCGGCCAGGCGCAGTGGCTCATGTCTGTAA
 TCCAGCACTTTGGGAGGTCAAGCTGGGCGGATCACTTGAGGTGAGGATTTGAGACCAG
 GCTGACCAACGTGGTGAAACCCCGTCTCTACGAAACATACAAAAATTAGTCGGGTGTGGT
 GGCACACTCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATAATTGCTTGAAT
 CTGGGAGGCGGAGGTTGCAGTGAGCCGAGATCATACCACTGCACTTCAGCCTGGGCAGCA

24764 GCCTGTGACTATCAGCTCTTTCAAGTCCTTCCTGGGGACCGGGGTGAGCAGGGATGGGTT
 GGAGCTCAGGATAGGGGGCAGCGGGGAGGCGAGCAGACTGACCACGCCCAGGATGGAGC
 CCAAGGTTACCCGGGTTACAGGGCTGTGAGGTGCTTCAGGCAGAGAGTAGGGTAAGAT
 AATCAGTGGAGGTAAGAGGACATAAAATACCTGTAACCCAACGATGTAGGGTCATGAGAT
 TGTCTTGGCTCAGTGTGAGAGAGAGGTACCAAAGGTCATCTTCCTAAAATTTAAAAGACA
 [A,G]
 TAAGATTGTCCAGGGTCCGGCCAGGCGCAGTGGCTCATGTCTGTAATCCCAGCACTTTGG
 GAGGTCAAGCTGGGCGGATCACTTGAGGTGAGGATTTGAGACCAGGCTGACCAACGTGG
 TGAAACCCCGTCTCTACGAAACATACAAAAATTAGTCGGGTGTGGTGGCACACTCCTGTA
 GTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATAATTGCTTGAATCTGGGAGGCGGAGG
 TTGCAGTGAGCCGAGATCATACCACTGCACTTCAGCCTGGGCAGCAGAGCGAGACTCTGT

24780 TCTTTCAAGTCCTTCCTGGGGACCGGGGTGAGCAGGGATGGGTTGGAGCTCAGGATAGGG
 GGCAGCGGGGAGGCGAGCAGACTGACCACGCCCAGGATGGAGCCCAAGGTTACCCGGGT
 TCACAGGGCTGTGAGGTGCTTCAGGCAGAGAGTAGGGTAAGATAATCAGTGGAGGTAAG

FIGURE 3S

AGGACATAAAATACCTGTAACCCAACGATGTAGGGTCATGAGATTGTCTTGGCTCAGTGT
 GAGAGAGAGGTACCAAAGGTCATCTTCTAAAATTTAAAAGACAATAAGATTGTCCAGGG
 [A,T]
 CCGGCCAGGCGCAGTGGCTCATGTCTGTAATCCCAGCACTTTGGGAGGTCAAGCTGGGCG
 GATCACTTGAGGTGAGGAGTTTGAGACCAGGCTGACCAACGTGGTGAAACCCCGTCTCTA
 CGAAACATACAAAAATTAGTCGGGTGTGGTGGCACACTCCTGTAGTCCCAGCTACTCAGG
 AGGCTGAGGCAGGAGAATAATTGCTTGAATCTGGGAGGCGGAGGTTGCAGTGAGCCGAGA
 TCATACCACTGCACTTCAGCCTGGGCAGCAGAGCGAGACTCTGTTTAAAAAAAAAAAAAA

24856 GCAGACTGACCACGCCCAAGGATGGAGCCCAAGGTTACCCGGGTTACAGGGCTGTGAGG
 TGCTTCAGGCAGAGAGTAGGGTAAGATAATCAGTGGAGGTAAGAGGACATAAAATACCT
 GTAACCCAACGATGTAGGGTCATGAGATTGTCTTGGCTCAGTGTGAGAGAGAGGTACCAA
 AGGTCATCTTCTAAAATTTAAAAGACAATAAGATTGTCCAGGGTCCGGCCAGGCGCAGT
 GGCTCATGTCTGTAATCCCAGCACTTTGGGAGGTCAAGCTGGGCGGATCACTTGAGGTCA
 [A,G]
 GAGTTTGAGACCAGGCTGACCAACGTGGTGAAACCCCGTCTCTACGAAACATACAAAAAT
 TAGTCGGGTGTGGTGGCACACTCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGA
 ATAATTGCTTGAATCTGGGAGGCGGAGGTTGCAGTGAGCCGAGATCATACCACTGCACTT
 CAGCCTGGGCAGCAGAGCGAGACTCTGTTTAAAAAAAAAAAAAAAAAAAAAGACTGTCCA
 CGGACAAGTGACAGAAGGGAGTGTTTCTGACCTCAATTTGTAGGATGGGCTGGGCATGG

26136 ATGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAAG
 TCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAGCCTCGTCTCTACTAAAAATACAA
 AAAATTAGCCAGGCATGGTGGCAGGCGCCTATAGTCCCAGCTACTCAGGAGGCTGAGGCA
 GAAGAATCACTTGAACCCAGGAGGTGAAGGTTGCAGTGAGCCAAGATTGCGCCACTGCAC
 TCCAGCCTGGCGACAGAGTGAGACTCCGTCTCAAAAAAAAAAGAAAAATAGATTGTCTAG
 [A,G]
 GTCGAGTGAGAGAAGGGAGTGTAGAAGTTTGTCTGATCTTAAGTTTGTAGCATCATGAGA
 TTGTTTCAGGCTCAACCTGATGGGATGGGAGACTAAAGGGCATCTGGGCTTAGATTTGTGA
 GAACTAAGTTTGTTCACCACTGGGACCCTGAAGTTATCTGAAGTTGGGACGGGAGAGAGG
 CAAATGGATAGCCGCGGAAGCATGAGATTGCTGTCTGACAGGGAGAAGCAAGGGATTG
 AGCGTATTCACGCTGAAGTACATGGCATGAGGTTGGCTGGATATTAGGAAAGGATGCTTG

26331 CCCAGGAGGTGAAGGTTGCAGTGAGCCAAGATTGCGCCACTGCACTCCAGCCTGGCGACA
 GAGTGAGACTCCGTCTCAAAAAAAAAAGAAAAATAGATTGTCTAGGGTCGAGTGAGAGAA
 GGGAGTGTAGAAGTTTGTCTGATCTTAAGTTTGTAGCATCATGAGATTGTTTCAGGCTCAA
 CCTGATGGGATGGGAGACTAAAGGGCATCTGGGCTTAGATTTGTGAGAACTAAGTTTGT
 CACCACTGGGACCCTGAAGTTATCTGAAGTTGGGACGGGAGAGAGGCAAATGGATAGCCG
 [C,G]
 GGAAGCATGAGATTGTCCTGTCTGACAGGGAGAAGCAAGGGATTGAGCGTATTCACGCTG
 AAGTACATGGCATGAGGTTGGCTGGATATTAGGAAAGGATGCTTGTGGTTGTTTCAGGTGT
 TGAGTGTGAGGCCACAAGCTCGTGCAGGCTGGAAGTGGGAAGTTATTCAAGTTTCATGGTG
 ACAGCAGCATGGGATTGGCTGGGAGTGTTGTGGGGAGGGGTAGGGTGAGCAGGAAGTT
 GTTTGGCGGGGGTGGTCTAGGGTGGTCTAAGTTTGCCCAAACTTTTACTGCAGGTTGTC

28362 CTGAGGCAGGAGAATTGCTTGAACCCGGAAGGTGGAGGTTGCAGTGAGTGGAATCACAC
 CACTGTACTCCAGCCTGGGTGACAGAGCAAGACCCTATCTCAAAAACAAACAAACAAACA
 AATGAACAAACAAAAATTTCTGAGTGTGGTGATATGAGACTGTAATCCTACCTACTTG
 GGAGGCTGAGCTGGGAGAATCACCAGAGCCCTGGGAGGTTGAAGCTGCAGTGAGCAGTGA
 CTGGGCCCTGCACTCCAACCTGGAGGACAGAGGGAGATCCTGTCTCAAAAACAAAAAA
 [A,G,C]
 TAAGAGCCCTAAGAAAGGTGTTGAGTCGGGTATGACACTCAACCCAGATGCCAGAGAGGA
 TCCTGTCTGGCCGGACACAGTGGCTCAGGAGGGTAATCCAGTACTTTGGGAGGCTGAGG
 TAAGAGGATTGCTTAAGTTCAGGAGTCCAGAGCAGCCTGGGCAATACAGTGAGATCACA
 TCTCACTAAATAAATAAATAAAGGATCCTATCACACAAAGAGGGTTTAGGACTTCCTTCC
 CCAACATTTTGGGGTGATATGCCTCTTTTCTACTGTATATATGGGAGAGTGACTAACTG

28745 GCTCAGGAGGGTAATCCAGTACTTTGGGAGGCTGAGGTAAGAGGATTGCTTAAGTTCAG
 GAGTCCGAGAGCAGCCTGGGCAATACAGTGAGATCACATCTCACTAAATAAATAAAATAAA

FIGURE 3T

GGATCCTATCACACAAAGAGGGTTTAGGACTTCCTTCCCCAACATTTTGGGGTGATATG
CCTCTTTTCTACTGTATATATGGGAGAGTGACTAACTGAAATTCCATCAGAATTAGAAAC
AAATAGCATCATTACCCATGAGTCAATAAGGGCTGTGAGGATGGGCCCTTTCATTGCCC
[T, C]
CACCTTCTTCCTCTTCCTGTACAGATAACCCATCTGTGCAAAGAAGAGAAAAAGAGGTT
GGGTGTGGTGGCTCACATCTGTAATCCCAGCACTTTGGGAGGCTAAGGTGGAAGGATTTT
GAGCCCAGGAGTTTGAGACCAGCCTGGGCAACATAGTGAGACCCCATTTCTACAAAAAAA
TACAAAGATTGGCCAGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCTGAG
GCAGGCGGATCATGAGGTCAGGAGATCGAGACCATCCTGGCTAGGTGAAACCCCGTCTCT

FIGURE 3U